NCEE 2012-4020

U. S. DEPARTMENT OF EDUCATION

Effects of the Kentucky Virtual Schools' hybrid program for algebra I on grade 9 student math achievement

NATIONAL CENTER FOR EDUCATION EVALUATION AND REGIONAL ASSISTANCE

Institute of Education Sciences

Effects of the Kentucky Virtual Schools' hybrid program for algebra I on grade 9 student math achievement
Final Report

April 2012

Authors:

Linda Cavalluzzo, Ph.D.
CNA Education

Deborah L. Lowther, Ph.D.
Center for Research in Educational Policy
University of Memphis

Christine Mokher, Ph.D.
CNA Education

Xitao Fan, Ph.D.
University of Virginia

Project Officer:
Karen Armstrong
Institute of Education Sciences

NCEE 2012–4020

U.S. Department of Education

U.S. Department of Education

Arne Duncan

Secretary

Institute of Education Sciences

John Q. Easton

Director

National Center for Education Evaluation and Regional Assistance

Rebecca A. Maynard

Commissioner

April 2012

This report was prepared for the National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, under contract ED-06CO-0021 with Regional Educational Laboratory Appalachia administered by CNA Education.

IES evaluation reports present objective information on the conditions of implementation and impacts of the programs being evaluated. IES evaluation reports do not include conclusions or recommendations or views with regard to actions policymakers or practitioners should take in light of the findings in the report.

This report is available on the Institute of Education Sciences website at http://ncee.ed.gov and the Regional Educational Laboratory Program website at http://edlabs.ed.gov.

Alternate Formats Upon request, this report is available in alternate formats, such as Braille, large print, audiotape, or computer diskette. For more information, please contact the Department's Alternate Format Center at 202-260-9895 or 202-205-8113.

Contents

Tables

Figures

Disclosure of potential conflict of interest

Regional Educational Laboratory Appalachia at CNA was the prime contractor for this evaluation. Center for Research in Educational Policy at the University of Memphis and the Collaborative for Teaching and Learning were subcontractors.[1] None of these organizations or their key staff members has financial interests that could be affected by findings from the study. None of the members of the Technical Working Group, convened by the research team to provide advice and guidance, has financial interests that could be affected by findings from the study.

[1] Contractors carrying out research and evaluation projects for IES frequently need to obtain expert advice and technical assistance from individuals and entities whose other professional work may not be entirely independent of or separable from the tasks they are carrying out for the IES contractor. Contractors endeavor not to put such individuals or entities in positions in which they could bias the analysis and reporting of results, and their potential conflicts of interest are disclosed.

Acknowledgements

This study was made possible by the collaboration and hard work of many individuals beyond the authors. The study team would like to thank the many districts, schools, and teachers who participated. Their cooperation and support are much appreciated. This study also would not have been possible without the Kentucky Department of Education, which supplied course seats and registration for the online courseware, technical support to participants, and data for the analysis.

Kelly Vosters at CNA and Todd Zoblotsky, Cindy Muzzi, and Rachael Best at the Center for Research in Educational Policy, University of Memphis, made important contributions to the data analysis and editing of the report. Jo Ann Mosier and Roland O’Daniel at the Collaborative for Teaching and Learning provided invaluable support to the participating schools and teachers through the professional development program. The research team would also like to thank the Technical Working Group for its suggestions and support. The members of the Technical Working Group are Johannes Bos, Margaret Burchinal, Laura Desimone, Barbara Goodson, and Samuel Stringfield.

Executive Summary

Background and motivation

This study, through a rigorous randomized controlled trial (RCT), quantifies the effectiveness of the Kentucky Virtual Schools' hybrid program for algebra I, an intervention combining online instruction with face-to-face classroom teaching to increase student learning and achievement in grade 9 algebra I, a course required by Kentucky for high school graduation. In this program, teachers engage in ongoing professional development—through online content used in face-to-face and distance settings—to learn how to teach an algebra I course structured on research-based strategies in a hybrid classroom.

The Kentucky Department of Education (KDE) has been involved with online education through Kentucky Virtual Schools (KVS, formerly Kentucky Virtual High School) since 2000, when it introduced online courses for high school students who, for various reasons, could not take a needed course at their regular school. Soon after, Kentucky Virtual Schools began offering teachers its online courseware—to aid instruction in face-to-face classrooms. The goal of these hybrid courses was to increase the expected achievement outcomes of students whose classroom teachers were inexperienced or less successful than desired in specific subject areas. The hybrid instructional approach was seen as "join[ing] the best features of in-class teaching with the best features of online learning to promote active independent learning" (Garnham and Kaleta 2002, para. 1). Although the Kentucky Department of Education did not systematically collect data on gains associated with hybrid instruction, education leaders in the department reported substantial improvement in student achievement for teachers who adopted the approach. Based on this perception, and demand from educators across the state, Kentucky Virtual Schools furnished courseware for hybrid courses in 26 classrooms, reaching over 500 students in school years 2005/06 and 2006/07.

A rigorous evaluation of the Kentucky Virtual Schools' program for algebra I was motivated by the needs and interests of educators and policymakers in Kentucky, the Appalachia Region, and the nation. The 2005 Regional Advisory Committee report for Appalachia (CNA 2005a) indicated that a focus on math instruction and the use of technology are important areas of educational need in the region. These needs align with a broader national interest in improving student achievement in math. As noted in a report by the National Research Council (1989):

> More than any other subject, mathematics filters students out of programs leading to scientific and professional careers. From high school through graduate school, the half-life of students in the mathematics pipeline is about one year; on average, we lose half the students from mathematics each year, although various requirements hold some students in class temporarily for an extra term or a year. Mathematics is the worst curricular villain in driving students to failure in school.

> When mathematics acts as a filter, it not only filters students out of careers, but frequently out of school itself (p. 7).

Literature review

Although the effectiveness of hybrid algebra programs has not been studied rigorously, three recent RCTs examining the effect of computer-assisted instruction in algebra I on student achievement are particularly relevant to the current study. First, Dynarski et al. (2007) conducted RCTs of 10 software products, including three for algebra I, and found no statistically significant differences in student achievement between the treatment group and control group. Next, Campuzano, Dynarski, Agodini, and Rall (2009) invited participants who used one of two chosen products in the first cohort of Dynarski et al. (2007) to continue participating for a second year and found a statistically significant positive effect size of 0.15 on student achievement for students of teachers who used the algebra products for a second year. Third, Barrow, Markman, and Rouse (2009) tested the impact of a computer-assisted instruction program for algebra I and reported a statistically significant overall intent-to-treat effect size of 0.17 and a 0.25 estimated effect size for the treated sample.

While these studies all looked at the impacts of computer-assisted instruction on algebra I, some of the programs used in them served as the core curriculum in a computer-only setting. The current study focuses solely on a hybrid—as opposed to computer-only—program, where teachers still have a central instructional role and online resources are a tool to enhance instruction and learning. Also, the study randomizes at the school level to eliminate a potential spillover effect between teachers in the same school.

Intervention and theory of action

The Kentucky Virtual Schools' hybrid program for algebra I seeks to boost student achievement and increase grade 10 math course-taking by improving classroom instruction. The program intends to achieve these goals by introducing into the classroom online courseware that provides lessons and exercises as part of regular instruction and by intensive and sustained professional development intended to help teachers integrate online resources with their lessons and improve their content knowledge and instructional practices.

According to the theory of action, change begins with professional development. The professional development experience for hybrid algebra I, which begins the summer before the school intervention year, is guided by instruction specialists and uses an online professional development program for algebra I teachers, as well as other resources. The summer program is intended to improve both the algebra I content knowledge and research-based instructional practices. It begins with orientation and instruction in a hybrid format. It then moves to a distance format, with teachers meeting online in guided weekly sessions designed to increase their conceptual understanding of math content and pedagogy. During the school year, monthly online sessions, also guided by instruction specialists, allow teachers to share their successes and challenges to improve their understanding of, and ability to implement, recommended instruction practices. Topics

for monthly discussions are selected by the instructional specialists and informed by site visits to schools, where the instruction specialists observe implementation and help guide improvement. The increased knowledge, discussions, and coaching are intended to help teachers use the new or unfamiliar classroom instruction materials and activities, including the hybrid algebra I courseware for individual student use and the professional development courseware for whole-group instruction. The resulting changes in classroom practices are expected to increase student achievement and math advancement.

The key Internet-based resource for students is the Kentucky Virtual Schools/Kentucky Department of Education-selected courseware, distributed by the National Repository of Online Courses. Kentucky Virtual Schools and the Kentucky Department of Education chose the repository because it satisfies their quality and content standards. Further, it is customizable by educators; and the unlimited-use license associated with Kentucky Virtual Schools' membership in the National Repository of Online Courses consortium makes program expansion affordable.[2] The last two features have particular appeal for going to scale because the content can be tailored to meet the unique needs of different school systems, and doing so likely would not be cost prohibitive. The main resource for the summer professional development is Spotlight on Algebra I, courseware focusing on content and pedagogy. Teachers continued to have access to this resource throughout the school year and could use it for instruction during face-to-face lessons. Spotlight on Algebra I was developed by the Southern Regional Education Board with financial support from the AT&T Foundation (Southern Regional Education Board 2001). The courseware is free to Kentucky through its membership with the board.

Math instruction specialists are also a key resource for teachers. The instructional specialists developed the year-long syllabus and a binder with resources for teachers. The instructional specialists led summer professional development lessons, held monthly school-year discussions, and visited each school site twice. Instructional specialists were furnished by the Collaborative for Teaching and Learning, a Louisville-based professional development provider.

In Kentucky Virtual Schools, instruction time is 60 percent face-to-face instruction and 40 percent student use of online resources. The hybrid algebra I program prescribes a standard three-part procedure for each lesson. Each lesson begins with an activity that *activates prior knowledge* associated with the lesson's learning goals. This can involve whole-class question-and-answer sessions, warm-up problems, discussions, or activities, as well as individual or group tasks completed by the student. The second procedure introduces *new learning*. This may include new vocabulary and face-to-face and online learning activities. Journal writing and other writing activities are also encouraged, to reinforce newly introduced algebra I concepts. The final procedure, *lesson closure*,

[2] Kentucky Virtual Schools provides the Blackboard online learning system, access to the repository of online courses and instructional resources, registration services, a 24/7 help desk, initial teacher training, and hosting of the online environment. To obtain access to Kentucky Virtual Schools' blended learning resources, schools or districts purchase accounts at $25 per student, which offer unlimited access to enrollment in online courses, community and instructional resources at no additional cost. Teacher and facilitator accounts are provided for free (see www.education.ky.gov/KYVSAssets/KYVS%20Blended%20Instruction.pdf).

provides students with an opportunity to reflect on what they learned during the lesson. Reflection helps students process and retain new information. Reflection strategies include note-taking, exit slips, student journals, and small-group discussion. Teachers also use reflection as a formative assessment tool.

Research questions and outcomes

Two primary confirmatory outcomes were evaluated. The first outcome measure was the score on the pre-algebra/algebra portion of the American College Testing (ACT) PLAN® assessment (PLAN), which measures students' math achievement in the fall of grade 10. The second outcome is an indicator of the students' math course enrollment in grade 10. The confirmatory analysis included two primary confirmatory research questions:

1. What is the impact of the Kentucky Virtual Schools' hybrid program for algebra I on math achievement levels in pre-algebra/algebra in the fall of grade 10?
2. What is the impact of the Kentucky Virtual Schools' hybrid program for algebra I on students' math course enrollment in grade 10?

The Kentucky Department of Education requires all students in the fall of grade 10 to take the PLAN. This assessment, part of the ACT testing program, evaluates whether students are on track for college readiness and is an important part of the department's state accountability program. A grade 9 intervention that improves college readiness in grade 10, after controlling for performance on the ACT given in grade 8 (the EXPLORE), would interest the Kentucky Department of Education. A limitation is that treatment was administered during one school year and the PLAN was conducted in the fall semester the following year. A meta-analysis on the effects of summer vacation on achievement test scores (Cooper, Nye, Charlton, Lindsay, and Greathouse 1996) found that test scores declined by about one-tenth of a standard deviation between the end of one school year and the start of the next, and the effect of summer break was more detrimental for math than for reading. Similarly, the impact of the hybrid program for algebra I may have diminished during the time lag between the administration of the treatment and the PLAN. However, both treatment and control students were subject to the effect of summer break, so the test scores for both groups of students are likely to have declined from the spring of the intervention year to the fall of the postintervention year.

Grade 10 math course enrollment is another important outcome. As Muller and Schiller (2000) note, both completion of a high school diploma and the number of math credits earned in high school are important because "failing to obtain either of these indicators is a hurdle to further educational or occupational opportunities, although at different levels." (p. 200) Further, the highest level of math completed in high school is one of the strongest predictors of whether a student will enter postsecondary education, be prepared for college-level courses without remediation, and complete a bachelor's degree (Adelman 1999, 2006; Long, Iatarola, and Conger 2009).

Exploratory analyses examined whether the impact of the hybrid program for algebra I was significantly different between male and female students, between students in cohort I and cohort II, and between students in rural and nonrural schools.[3]

Study design

This study examined Kentucky public schools with grade 9 algebra I classes. A volunteer sample of 47 schools (30 of which are in rural areas) was randomly assigned to the treatment and control conditions. The intervention was applied in one school year and evaluated the next fall.

A power analysis was conducted under a range of assumptions to determine the appropriate sample size. Allowing for potential attrition of 20 percent at the school level, the power analysis indicated that a sample of at least 47 schools, with approximately 6,000 students was needed. In the actual sample, there was no attrition at the school level, meaning that the confirmatory tests had greater than 80 percent power to detect true impacts of .20 standard deviations or larger.[4]

The study enrolled 25 schools in the 2007/08 school-year (13 treatment and 12 control) and 22 schools in 2008/09 (11 and 11). Randomization occurred at the school level. All algebra I teachers and students in a school were assigned to the school's treatment condition, and all algebra I teachers in a treatment school were asked to use the intervention. School-level randomization was chosen, as opposed to within-school randomization, because it minimizes the potential for spillover effects from the treatment to control classrooms, which could lead to underestimating the impacts of the hybrid program.

In addition to the formal statistical analysis of grade 10 achievement and grade 10 math course enrollment, information was collected from teacher surveys and classroom observations describing classroom activities. This information was used to describe the extent to which the treatment was implemented with fidelity, as well as to compare the classroom activities in treatment and control classrooms.

Analysis and findings

The intent-to-treat sample consisted of all grade 9 students in treatment and control schools enrolled on September 1 in a course culminating in the completion of algebra I credit, regardless of whether the intervention was used as intended. This sample comprised 6,908 students, 61.4 percent of whom attended rural schools. Analyses for the

---

[3] To increase sample size, all the control schools in cohort I were asked to enter cohort II. Six accepted the invitation. The six schools were separately randomized into cohort II, with three entering the control group and three the treatment group. The remaining 16 of the 22 schools in cohort II were new to the study. Only control schools from cohort I were rerandomized to avoid treatment-to-control spillover across cohorts and maturation effects from schools participating in the treatment for a second year. We had planned to examine whether the impact of the hybrid program differed among students with different levels of pretest performance or among schools in different strata (schools initially assigned to cohort I, schools from the cohort I control group that were rerandomized into cohort II, or new schools randomized into cohort II). However, there was insufficient power to justify conducting these analyses.

[4] Although five schools withdrew from the study, data for these schools were collected and included in the analysis.

confirmatory and exploratory research questions on grade 10 math enrollment were conducted using data from students in the intent-to-treat sample who were enrolled in a Kentucky public school during the postintervention year, while analyses for the PLAN outcome were further restricted to students who also had PLAN test scores.[5] The analyses used two-level hierarchical linear models to nest students within schools and assess differences in outcomes between the treatment and control schools.

The findings indicate that the treatment has no statistically significant effect for either outcome. Students in the treatment group did not achieve scores on the pre-algebra/algebra portion of the PLAN that were significantly different (higher or lower) than did students in the control group. Nor were students in the treatment group significantly more or less likely than students in the control group to enroll in a math course above algebra I the year after the intervention. Sensitivity analyses were conducted to help establish the robustness of the impact estimates, but none produced a change from the results of the confirmatory impact analysis.

Exploratory analyses were conducted to determine whether the impact of the hybrid program differed by student gender, student cohort, or school rural status. No statistically significant differences were observed for these study subgroups for the impacts of the intervention on math achievement in pre-algebra/algebra on the PLAN or on grade 10 math course enrollment.)

Study limitations

There are limitations due to the fidelity of implementation as 4 of the 24 treatment schools were noncompliant, which means they did not participate in the intervention. In addition, 20 of the 63 teachers did not participate in any component of the intervention; 6 because their school withdrew from the study after randomization; 14 others neither attended the professional development sessions nor used the online student courseware. Twelve of those teachers were never reported by participating schools as teaching algebra I, and were not identified by the research team until student enrollment records were gathered from the Kentucky Department of Education late in the school year. Less than 50 percent of the treatment sample (47.6 percent of teachers in the summer and 42.9 percent of teachers in the school-year) had high or moderate attendance for the professional development components. The percent of teachers rated as having low engagement during the professional development sessions was 19 percent in the summer and 30 percent in the academic year. Further, 65 percent of treatment students had ratings of no or low use of the Kentucky Virtual School's online algebra I materials during the period for which use data were collected. The data available for fidelity of implementation included the number of student connections per week from the electronic archives of the courseware, but not the actual amount of class time spent using the courseware so this is only an approximation of whether the online materials were used as intended.

[5] In both the treatment and control groups, 93 percent of students in the sample had outcome data for grade 10 math course enrollment. Response rates for the PLAN were 86 percent for the control group and 84 percent for the treatment group.

This study also has several limitations relating to the generalizability of the findings to other settings and contexts. One limitation is that the sample schools volunteered to participate. These schools may differ from the broader population of Kentucky high schools in both observable and nonobservable characteristics. As a result, the findings are not generalizable beyond the sample. Study results also cannot be generalized to other models of hybrid instruction, or to the current model of hybrid instruction implemented with different professional development or student courseware.

The student courseware was the latest version available from the National Repository of Online Courses when the intervention began. However, the Virtual School mistakenly installed an early version of the student courseware. The older courseware had not been through careful review, and student exercises had many errors. Complaints from teachers to instructional specialists led to an inquiry by the research team, the discovery of the source of the problem, and the problem's resolution. Updated student courseware was installed when the second semester of the first cohort began and was used for the rest of the study. It could be expected that such an error would negatively affect continued use of the intervention, but the data showed no evidence of this. Rates of participation in professional development, and frequency of student logins to the courseware were similar for cohorts I and II. There were also no statistically significant differences between the impacts of the intervention by cohort on math achievement in pre-algebra/algebra on the PLAN or on grade 10 math course enrollment.

Other limitations pertain to the students. Data were missing on the PLAN outcome for students who were not promoted, left the Kentucky public school system, or missed testing in grade 10. The response rate for the PLAN was 86 percent for the control group and 84 percent for the treatment group, meaning that the results of the analysis of student math achievement apply only to students who were enrolled in algebra I at a participating school in grade 9 and promoted to grade 10. They do not generalize to algebra I students in participating schools who were not promoted to grade 10.

Finally, the results of the study apply only to grade 9 students enrolled in courses leading to algebra I credit. The results cannot be generalized to algebra I courses that do not fulfill the algebra I credit requirement, such as algebra I part A (the first course in a two-year sequence) or algebra I lab (an elective course). Further, the results do not apply to online courses for other subjects or grade levels provided through Kentucky Virtual Schools.

1. Study Overview

This study, through a rigorous randomized controlled trial (RCT), quantifies the effectiveness of the Kentucky Virtual Schools' hybrid program for algebra I, an intervention combining online and face-to-face instruction with the goal of increasing learning and achievement in grade 9 algebra I, a course required by Kentucky for high school graduation. In this program, teachers engage in ongoing professional development designed to improve their algebra I content knowledge and train them in research-based hybrid instructional practices. The hybrid professional development familiarizes teachers with the environment in which their students will learn. In face-to-face professional development, instructional specialists model algebra I instructional practices and engage teachers in hands-on learning, designed to increase teacher content and pedagogical knowledge. Online sessions are intended to give teachers direct experience learning with digital tools in an effort to prepare them to integrate the online algebra I courseware with their instruction.

This study used a two-cohort sample to reach a size sufficient for hypothesis testing. With 25 high schools in year 1 (cohort I: 13 treatment and 12 control) and 22 in year 2 (cohort II: 11 and 11), the randomized sample included 6,908 students, 61.4 percent in rural schools.[6] The study used two-level hierarchical models to nest students within schools and assess differences in outcomes between the treatment and control schools. There were two primary confirmatory outcomes. The first outcome measure was the score on the pre-algebra/algebra portion of the American College Testing (ACT) PLAN® assessment (PLAN), which measures the impact of the intervention on grade 9 students' math achievement in the fall of grade 10. The second outcome assessed the intervention's impact on grade 9 students' grade 10 math course-taking.

Background and policy context

Online courses, distinct from hybrid courses, provide computer-based instruction to students geographically separated from their teacher. Students may access course materials wherever they can access the Internet. Often, students access their online course(s) from home or a school computer lab. Some online courses include synchronous sessions, enabling real-time communication between the class members and the online teacher; others are fully asynchronous—the courses include no real-time communication.

In hybrid courses, students engage in a variety of face-to-face whole-class, small-group, and individual activities to learn the targeted knowledge and skills of the courses, just as they would in regular courses. However, hybrid courses also incorporate online instructional materials that include self-paced tutorials and activities, which provide feedback based on student performance. In addition, the hybrid format provides

[6] To increase sample size, all the control schools in cohort I were asked to enter cohort II. Six accepted the invitation. The six schools were separately randomized into cohort II, with three entering the control group and three the treatment group. The remaining 16 of the 22 schools in cohort II were new to the study. Only control schools from cohort I were rerandomized to avoid treatment-to-control spillover across cohorts and maturation effects from schools participating in the treatment for a second year. Sensitivity analyses indicated that excluding the rerandomized schools did not affect the findings.

classroom teachers and classroom facilitators opportunities to work with individual students during the class time devoted to individual student use of the online materials.

Hybrid instruction (also called blended instruction) has many formats. Louisiana's Algebra I Project couples uncertified teachers with certified teachers. The uncertified teacher is in the classroom with students; the certified teacher supports instruction from a distance. The formats also vary by the amount of time devoted to students' use of online resources. For the Kentucky Virtual Schools' hybrid program for algebra I, teachers and students meet in their regularly assigned classroom or a computer lab. Students are expected to have individual access to online course materials at least two days a week (40 percent of class time). Classroom teachers work individually with students to support learning during these online sessions, and they may gather students for a mini-lesson during this online period if a common problem needs to be addressed. In the Kentucky model, classroom teachers are also expected to participate in a professional development program to help them improve their algebra I instructional practices and effectively integrate the online materials with their instruction.

The Kentucky Department of Education (KDE) has been involved with online education through Kentucky Virtual Schools (KVS), formerly Kentucky Virtual High School) since 2000, when it introduced online courses for high school students who, for various reasons, could not take a needed course at their regular school. Soon after, Kentucky Virtual Schools began offering teachers its online courseware to aid their instruction in face-to-face classrooms. The goal of these hybrid courses was to increase the expected achievement outcomes of students whose classroom teachers were inexperienced or less successful than desired in specific subject areas. The hybrid instructional approach was seen as "join[ing] the best features of in-class teaching with the best features of online learning to promote active independent learning" (Garnham and Kaleta 2002, para. 1). Although the Kentucky Department of Education did not systematically collect data on gains associated with hybrid instruction, education leaders in the department reported substantial improvement in student achievement for teachers who adopted the approach. Based on this perception, and demand from educators from across the state, Kentucky Virtual Schools piloted hybrid courses, furnishing courseware to 26 classrooms, reaching over 500 students in school years 2005/06 and 2006/07.

The Kentucky Department of Education is a member of the Southern Regional Education Board, which provides an unlimited use license for its professional development courseware, and the National Repository for Online Courses, which provides an unlimited use license to KDE for its student courseware. These and other online resources are made available through Kentucky Virtual Schools. For hybrid courses, Kentucky Virtual Schools charges $25 per student to cover course registration and access to the Blackboard platform, which houses the courseware. Blackboard also provides an online course site that facilitates two-way communication between teachers and students. It includes tools for social learning (discussion boards), monitoring (usage statistics), posting course materials (syllabus), and course planning (teacher lesson plan modules).

A rigorous evaluation of the Kentucky Virtual Schools' program for algebra I was motivated by the needs and interests of educators and policymakers in Kentucky, the Appalachia region, and the nation. The 2005 Regional Advisory Committee report for

Appalachia (CNA 2005a) indicates that a focus on math instruction and the use of technology are important areas of educational need in the region. These needs align with a broader national interest in improving student achievement in math. As noted in a report by the National Research Council (1989):

> More than any other subject, mathematics filters students out of programs leading to scientific and professional careers. From high school through graduate school, the half-life of students in the mathematics pipeline is about one year; on average, we lose half the students from mathematics each year, although various requirements hold some students in class temporarily for an extra term or a year. Mathematics is the worst curricular villain in driving students to failure in school. When mathematics acts as a filter, it not only filters students out of careers, but frequently out of school itself (p. 7).

Students in Kentucky must take four years of math, including algebra I, geometry, and algebra II, to graduate from high school; students who fail algebra I must retake the class before progressing to a subsequent math course, which hinders their taking an additional advanced math course and “ensure readiness for postsecondary education or the workforce based on the student’s individual learning plan” (Kentucky Department of Education 2006b, p. 1).

The Regional Advisory Committee report for Appalachia (CNA 2005a) further cites two of the top five areas of need in the region as identifying evidence-based curricula/programs and improving teacher quality. Improving teacher quality also is a national concern (CNA 2005b). Improving teaching in rural areas poses special challenges. Rural areas are often economically depressed and geographically and socially isolated, so attracting teaching candidates from more urbanized areas is difficult (McClure, Redfield, and Hammer 2003), limiting the pool of qualified candidates. These conditions can lead to low-performing rural schools with a limited supply of teachers, many of whom may have received much of their own education in the same low-performing schools (Loeb and Reininger 2004; Monk 2007; Strauss 1999).

Literature review

There is a considerable body of literature on both technology’s effectiveness in improving instruction and how technology can be used most effectively. However, the evidence on the effectiveness of technology in improving student performance is mixed. Kirkpatrick and Cuban (1998) distinguished three uses for computers in instruction. The simplest computer-assisted instruction uses software programs to provide tutorials and exercises. More sophisticated programs evaluate student knowledge and skills, skip content students have mastered, and move students to topics they have yet to master. Kirkpatrick and Cuban called this use computer-managed instruction. In a third use, computer-enhanced instruction, teachers use the Internet, or spreadsheets or presentation programs, to make projects or assignments more interesting or relevant to students and help ensure that students develop skills in using technology. The Kentucky Virtual

Schools' hybrid program in algebra I integrates algebra I courseware with regular classroom instruction, a use of technology best categorized as computer-assisted instruction, but in a distinctive hybrid form.

James Kulik has conducted a number of formal meta-analytic studies of experimental and quasi-experimental evaluations of the impacts of computer-assisted instruction on student outcomes. His 2003 study included 61 peer reviewed papers and dissertations published between 1990 and 2000 that reported on experimental or quasi-experimental evaluations of classroom technologies. Kulik sorted the studies into six categories according to the specific type of technology used and reported results for each category. A review of 16 studies of Integrated Learning Systems (such systems provide vertically aligned computer-assisted instruction over several grade levels) from the 1990s, included seven studies in math conducted for one half to one school year in grades 2-8. Kulik reported an average effect size of 0.38 for increased math test scores for those studies. Kulik also examined six studies of the effects of computer tutorial programs (modules that focus on a single topic) in social studies and science conducted for ten days up to six weeks in grades 3–12. The analysis yielded an average effect size of 0.36 for student test scores.

Murphy, Penuel, Means, Korbak, and Whaley (2001) reported an average estimated effect size of 0.45 for student achievement on math tests, based on 31 experimental or quasi-experimental studies of discrete educational software described in peer-reviewed papers, dissertations, and reports published by independent evaluators. The studies covered a range of grades (7 from PreK or Kindergarten, 14 from grades 1-5, 7 from grades 6-8, and 3 from high school) and had a median sample size of 96. The authors concluded that the study provided "evidence of a positive association between student achievement and the use of discrete educational software products to support instruction in math" (p. 38).

Similarly, Waxman, Lin, and Michko (2003) conducted a meta-analysis that included 42 studies of the effects of computer-assisted instruction and online activities for students in conventional education settings. The authors included studies employing experimental and quasi-experimental designs in K-12 classrooms and published between 1997 and 2003. Included studies covered a range of grades (40 percent from grades K-5, 40 percent from grades 6-8, and 20 percent from high school) and had a mean sample size of 184. The authors reported a weighted mean effect size of 0.45 for cognitive outcomes, which indicated a "small, positive, significant effect on student outcomes when compared to traditional instruction" (p. 11).

Whereas the previous studies evaluated computer-assisted instruction interventions in general, three recent RCTs examining the effect of computer-based algebra I instruction on student achievement are particularly relevant to the current study: Dynarski et al. (2007); Campuzano et al. (2009); and Barrow, Markman, and Rouse (2009). Dynarski et al. (2007) conducted RCTs of 10 software products, including 3 for algebra I, used in classroom instruction. One product served as the core curriculum; the other two supplemented the curriculum in a hybrid setting. Dynarski et al. (2007) recruited schools and districts with no experience with the products and randomized classrooms to treatment condition, found no statistically significant differences in student achievement between the treatment and control groups.

Campuzano et al. (2009) conducted further analysis on two of the products studied by Dynarski et al. (2007): one that supplements curriculum in a hybrid setting and one that serves as the core curriculum). They invited participants who used either product in the first cohort of Dynarski et al. (2007) to continue their participation for a second year. About 27 percent of the teachers volunteered to remain for a second year, and they retained their year 1 treatment or control status (using the same product) while teaching a new cohort of students. This approach allowed the researchers to estimate the effect of using the courseware for two years for teachers who chose to continue using the intervention. The algebra I analysis sample for this portion of the study included 24 teachers and 1,051 students with both posttest and pretest (actual or imputed) scores. Campuzano et al. (2009) reported a statistically significant positive effect size of 0.15 on student achievement for teachers who used the algebra products a second year. By contrast, the impact was not statistically different from zero for teachers who used the product for only one year.

Barrow, Markman, and Rouse (2009) tested the impact of a computer-assisted instruction program for algebra I in 142 randomly assigned classrooms from three urban districts (1,605 students). The program was designed so that the teacher's role in the classroom was to assist students when they needed additional help. Classes randomized into the treatment group met in a computer lab where students used the online curriculum at their own pace. The researchers reported a statistically significant overall intent-to-treat effect size of 0.17 and a 0.25 estimated effect size for the treated sample.

Overall, the literature reviewed here indicates that the impact of computer assisted instruction programs vary by product, subject, grade level, and level of implementation support. The meta-analyses reviewed in this section show that most studies that met inclusion standards found positive and, in some cases, large positive effects associated with computer tutorial and educational software products. However, notably, the more rigorous randomized controlled trials generally show less consistent evidence of positive effects. For instance, Dynarski (2007) found no effects for 16 different math and reading products on students in first, fourth, and sixth grades and in algebra I. Using a subset of the teachers and 10 products from the Dynarski et al. sample to study a second year of implementation, Campuzano et al (2009) found significant overall effects on achievement. Yet when the authors tested the impact of each of the 10 software products separately, only one product had a significant positive effect on test scores while nine products did not have any statistically significant effects. Furthermore, Campuzano et al (2009) note:

> Characteristics of districts and schools that volunteered to implement the products differ, and these differences may relate to product effects in important ways. The findings do not adjust for differences in schools and districts that go beyond measured characteristics but may be related to outcomes.

This means the findings may not be generalizable to a larger population of schools that do not volunteer to use the intervention. The evidence is also mixed on the whether the number of years of experience with computerized instruction on algebra I instruction influences the impact of the intervention on student outcomes. On the one hand, Dynarsky et al (2007) found no effects attributable to two algebra I software products, while Campuzano et al (2009) found a small effect (effect size=0.15) for teachers using

the same products in a second school year. On the other hand, Barrow, Markman, and Rouse (2009) found larger, statistically significant impacts among teachers using a different product for only one year.

While these studies looked at computer-assisted instruction impacts on algebra I, some programs used in them served as the core curriculum in a computer-only setting. This study focused solely on a hybrid—as opposed to computer-only—program, where teachers still had a face-to-face instructional role for a portion of the class time, rather than serving only as a facilitator. Also, the study randomized at the school level, eliminating the potential spillover effect between teachers in the same school.

The hybrid program for algebra 1 intervention

The hybrid algebra I intervention is structured to increase the algebra I knowledge and skills of grade 9 students through combining research-based face-to-face practices with interactive online instruction. For the face-to-face portion of the intervention, teachers implement instructional practices aligned with guidance from National Council of Teachers of Mathematics (NCTM 1989, 1991, 2000), North American Council for Online Learning (NACOL 2007), National Education Technology Standards (International Society for Technology in Education 1998, 2000), and Kentucky algebra I standards (Kentucky Department of Education 2006b), such as asking *why* and *what if* questions and using multiple representations of the same concept, such as number lines, graphs, diagrams, and computers, to explain algebraic processes and algorithms. Student activities include working in groups, writing, and talking, and using algebraic manipulatives, graphing calculators, and computers.

For the online portion of the intervention, students complete algebra I activities that provide self-paced, interactive instruction and immediate feedback to guide student learning. The classroom teacher prepares students for what they will learn during the online sessions and reviews learning goals, provides assistance to individual students or student groups as needed during the online sessions, and uses an end-of-session activity such as exit slips[7] to end the online session.

A key component of the hybrid algebra I intervention is the professional development program, designed to prepare teachers to implement a hybrid course that integrates face-to-face instruction with student use of online algebra I activities. The ongoing professional development provides training on hybrid algebra I instructional strategies, such as teacher modeling to reinforce student problem solving, reasoning, and communication through reading, writing, talking, and technology and manipulatives to acquire knowledge and skills.

Professional development is provided by two instructional specialists in mathematics from the Collaborative for Teaching and Learning (CTL), a Louisville-based professional development provider. The instructional specialists are experienced professional development providers with backgrounds in mathematics instruction at the secondary

[7] The exit slip (Billmeyer and Barton 1998) is a writing-to-learn strategy that attempts to bridge students into and out of new learning. At the end of a learning session, students write their responses to questions from the teacher that focus on synthesizing or summarizing new learning, making connections to the real-world context, or creating a bridge between new learning and learning to come during the next class period.

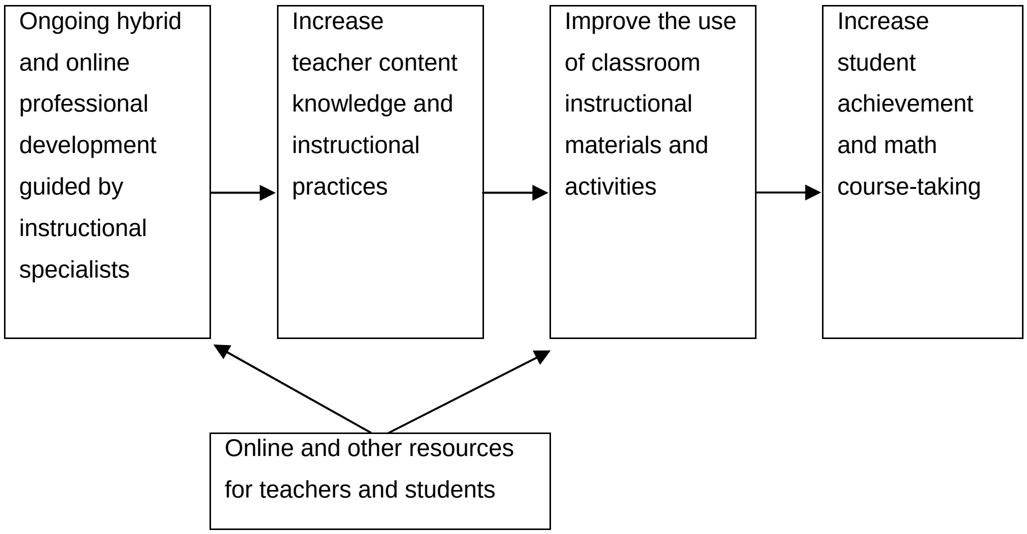
level and are trained by Kentucky Virtual Schools on the use of the primary online resources. The instructional specialists conduct the summer professional development sessions and offer constructive feedback to teachers during monthly online sessions during the school year, as well as during two on-site observations during the school year, on areas needing improvement. Algebra I teachers in schools randomly assigned as controls teach algebra I in a "business as usual" manner—using their normal strategies and resources.

Theory of action

The algebra I hybrid program seeks to increase student achievement and grade 10 math course-taking by improving classroom instruction. The program intends to achieve these goals by introducing into the classroom online courseware that provides lessons and exercises as part of regular instruction and by an intensive and sustained professional development experience.

Change begins with professional development (figure 1.1), which begins in the summer before the school intervention year. The summer program, designed to improve both the algebra I content knowledge and research-based instructional practices of teachers, is guided by instructional specialists and uses online professional development, as well as other resources. It begins with orientation and instruction in a hybrid format. It then moves to a distance format, with teachers meeting online in guided weekly sessions designed to increase their conceptual understanding of math content and pedagogy. During the school year, monthly online sessions, also guided by the instructional specialists, allow teachers to share their successes and challenges. Topics for monthly discussions are selected by the instructional specialists, informed by site visits to schools where they observe implementation and help guide improvement. The increased knowledge, discussions, and coaching are hypothesized to help teachers use the new or unfamiliar classroom instruction materials and activities effectively, including the hybrid algebra I courseware for individual student use and the professional development courseware for whole-group instruction. The resulting changes in classroom practices are expected to increase student achievement.

Figure 1.1. Hybrid algebra I theory of action



Research questions and overview of methods

The study addresses two confirmatory research questions:

1. What is the effect of the Kentucky Virtual Schools' algebra I hybrid program on grade 9 student math achievement measured in the fall of grade 10 (using the PLAN assessment)?

2. What is the effect of the Kentucky Virtual Schools' algebra I hybrid program on grade 9 students' math course enrollment in grade 10?

The Kentucky Department of Education requires all students to take the PLAN in the fall of grade 10. This assessment is part of the ACT testing program (ACT 2009), which evaluates whether students are on track for college readiness and is an important part of the department's state accountability program. A grade 9 intervention that improves college readiness in grade 10, after controlling for performance on the ACT assessment given in grade 8 (the EXPLORE), would interest the department.

Grade 10 math course enrollment is another outcome of student attainment that is separate from such achievement measures as test scores that measure mastery of curricular materials. As Muller and Schiller (2000) note, both completing a high school diploma and the number of math credits earned in high school are outcomes of education attainment, because "failing to obtain either of these indicators is a hurdle to further educational or occupational opportunities, although at different levels" (p. 200). In Kentucky, beginning with the graduating class of 2012, students must complete four years of math, including courses in algebra I, geometry, and algebra II.[8] This means that students who fail to progress to geometry or algebra II in grade 10 are at greater risk of failing to graduate from high school. Further, the highest level of math completed in high school is one of the strongest predictors of whether a student will enter postsecondary

[8] http://www.education.ky.gov/KDE/Instructional+Resources/Secondary+and+Virtual+Learning/High+School/Minimum+High+School+Graduation+Requirements.htm.

education, be prepared for college-level courses without remediation, and complete a bachelor's degree (Adelman 1999, 2006; Long, Iatarola, and Conger 2009).

Complementary exploratory analyses examined whether there are differences in the impact of the hybrid program between subgroups of students and schools.[9] The questions addressed were:

Exploratory research questions—student subgroup effects

1. Is the impact of the Kentucky Virtual Schools' hybrid program for algebra I significantly different for males than for females on:

    A. Students' math achievement in fall of grade 10, as measured by the PLAN assessment?

    B. Students' math course enrollment in grade 10?

Males and females have different math problem-solving strategies and learning styles (Carr and Davis 2001; Friedman 1995; Geary, Saults, Liu, and Hoard 1999). These gender differences may affect how students responded to the intervention.

2. Is the impact of the Kentucky Virtual Schools' hybrid program for algebra I significantly different for students in cohorts I and II on:

    A. Students' math achievement in fall of grade 10, as measured by the PLAN assessment?

    B. Students' math course enrollment in grade 10?

There are two notable factors that might have led to differences in the impacts of the intervention for cohorts I and II. One is the fact that during the first year of the intervention, the Virtual Schools installed an early version of the student courseware that had multiple errors in the lessons and assignments. These errors were not fully corrected until the second semester. A second is the fact that, in the first year, the instructional specialists lacked experience with the hybrid algebra program and this may have influenced the effectiveness of professional development.

[9] The combined recruiting sample consisted of three strata from the random assignment process: 25 schools initially assigned to cohort I, 6 schools from the cohort I control group that were rerandomized into cohort II, and 16 new schools randomized into cohort II. Whether the effect of the hybrid program differed by level of student pretest performance or among schools in different strata was intended for examination, but there was insufficient power to justify conducting these analyses.

Exploratory research questions—school subgroup effects

3. Is the impact of the Kentucky Virtual Schools' hybrid program for algebra I significantly different for students in rural and nonrural schools on:

   A. Students' math achievement in fall of grade 10, as measured by the PLAN assessment?

   B. Students' math course enrollment in grade 10?

Rural areas are often economically depressed and geographically and socially isolated, which makes it difficult to attract teaching candidates from more urban areas (McClure, Redfield, and Hammer 2003), limiting the pool of qualified candidates. Thus, the effects of the intervention may have differed in rural and nonrural schools due to the different characteristics of teachers and students in the two environments.

Study design overview

This study included two cohorts of schools, 25 schools in year 1 (13 treatment and 12 control) and 22 in year 2 (11 and 11). The study sample consisted of all grade 9 students who were enrolled in treatment and control schools on September 1st and were enrolled in a course culminating in the completion of algebra I credit. This included 6,908 students, 61.4 percent of whom attended rural schools. The analysis sample for the confirmatory and exploratory research questions on grade 10 math enrollment included the subset of students enrolled in any Kentucky public school during the postintervention year, while the analysis sample for tenth grade math achievement outcomes included the subset of students with 10th grade PLAN test scores.

Impacts were estimated using two-level hierarchical models to take account of the fact that students were nested within schools. No adjustments were made for multiple comparisons because math achievement and math course-taking are not the same outcome domain. Sensitivity analyses were conducted to help establish the robustness of the impact estimates.

Structure of the report

The rest of this report is divided into six chapters. Chapter 2 explains the study design, including random assignment, recruitment strategies, attrition and noncompliance, and the study sample. Chapter 3 describes the data and methods for the confirmatory and exploratory impact analyses. Chapter 4 provides details on the implementation of the intervention, including a descriptive report based on teacher surveys and classroom observations. Chapter 5 presents the confirmatory analysis impact findings, and chapter 6 provides the results from the exploratory analysis. Chapter 7 summarizes the findings and the study limitations.

2. Study Design

This study used a volunteer sample of schools, which were randomly assigned to the treatment and control groups. The intervention was applied in one school year and impacts were measured using data collected the following fall. A power analysis indicated that it was necessary to include approximately 50 schools and 6,000 students in the study sample if there was a school-level attrition rate of 20 percent in order to have 80 percent power to detect true impacts of .20 standard deviations or larger. (See appendix A for more details about how the power analysis was conducted.) In the randomization sample for this study, there was no attrition at the school level, meaning that the confirmatory tests were powered at greater than 80 percent.[10]In order to obtain a sufficient sample size, two cohorts of schools were recruited—one that used the intervention in 2007/08 and another that used it in 2008/09. Table 2.1 shows a timeline of key milestones.

Table 2.1. Timeline of key milestones for the study design, by cohort

Key milestones	Date for cohort I	Date for cohort II
Completed recruiting at the district and school levels	May 2007	May 2008
Completed memorandum of understanding with districts and schools	May 2007	May 2008
Conducted random assignment of schools	May 2007	May 2008
Began intervention implementation	May 2007	June 2008
Conducted classroom observations and teacher surveys[11]	March–April 2008	November 2008, February–March 2009
Completed intervention implementation	May 2008	May 2009
PLAN posttest administered	September 2008	September 2009
Follow-up data collection for grade 10 math enrollment	April 2009	February 2010

Sample recruitment

In order to identify schools eligible for this study, information from the Common Core of Data (http://nces.ed.gov/ccd/) were merged with data from the Kentucky Department of Education website on school performance in math. Among the 394 Kentucky schools with a grade 9, there were 237 regular schools that do not primarily provide special education, career and technical education, or alternative education. The resulting data file was used to select for possible inclusion in the study all regular schools in Kentucky in which algebra I is taught in grade 9 (N=216 schools). We also limited the recruitment to those schools in which a maximum of 60 percent of students were

[11] In the first semester of this project, the Office of Management and Budget did not provide clearance to begin data collection, so classroom observations could not be conducted until the spring semester for cohort I.

proficient in math, as indicated by either the Comprehensive Test of Basic Skills, version 5 for grade 9 students or the Kentucky Core Content Test (KCCT) for students in grade 8 or 11 (N=205 schools). This performance criterion targeted low-performing schools, which are more likely to benefit from an intervention designed to improve math achievement. After examining the distribution of test scores across schools, a natural break was found in the data at 60 percent proficient.[12] This selection process resulted in an initial list of 205 potentially eligible schools[13] that were sent a letter of invitation to participate in the study.

School recruitment entailed several steps. Information sessions were offered at the fall 2006 and 2007 conferences of the Kentucky Council of Teachers of Mathematics and at the spring 2007 and 2008 Kentucky Teaching and Learning Conferences to develop awareness about the RCT. These conferences are widely attended by Kentucky math teachers. Also, the study team sent a separate mailing with application packages to the building administrators of the 205 eligible schools. Each package included a cover letter signed by the Chief State School Officer, informing the administrators of their school's eligibility and encouraging them to participate. The packages were mailed in a large Kentucky Department of Education envelope with a brochure, an application describing the intervention, and a stamped, addressed return envelope. Once the packages were mailed, follow-up calls were made to district leaders and school principals to answer questions and further encourage participation. For cohort II recruiting, new packages were mailed and a new round of calls was made. The package for cohort II recruiting added a DVD with information about the program, including a discussion among participating teachers from cohort I on their experiences using the intervention.

Members of the study team gave presentations at each of seven spring meetings of regional cooperatives attended by superintendents from all districts in each region. This forum was used to familiarize superintendents with the intervention and the RCT. Follow-up meetings for principals and teachers expressing interest in the program were conducted after each cooperative meeting.

For cohort II, recruiting strategies were expanded. Kentucky Virtual Schools' administrators called schools with which they had a personal contact and reason to believe they would be interested in the intervention. Highly Skilled Educators[14] were informed of the program so that they could discuss it with teachers and building

---

[12] Most Kentucky schools had math proficiency rates less than 60 percent, so only 11 of 216 schools were eliminated due to this criterion.

[13] The list of 205 schools omits those in Jefferson County, a large urban/suburban area (Louisville), because the district was moving to a curriculum requiring algebra I in grade 8, and its schools, by criterion 1, were not eligible to participate. By limiting participation to algebra I students in grade 9, a more homogeneous sample of participants was created, which should reduce random variation in math performance.

[14] Highly Skilled Educators are teachers and administrators selected by the state to assist low-performing schools. To apply for a Highly Skilled Educators position, candidates must have Kentucky educator certification, a minimum of five years' teaching experience, involvement in teaching or administration in the last five years, and current full-time employment with a Kentucky school district. Selection involves performance assessments, including a written assessment, a simulated Highly Skilled Educators experience, delivery of professional development, a technology assessment, reference checks, and a Kentucky Department of Education site visit, where a department representative shadows the applicant and interviews the applicant and others at the school (Kentucky Department of Education 2010).

administrators with whom they worked. The study team also contacted the Dataseam Initiative, a nonprofit organization that provides classroom sets of Apple computers at little or no cost to qualifying schools in eligible Kentucky districts.

Schools assigned to the treatment group received the intervention, including professional development and materials for all participating teachers and follow-up support throughout the year. The cost of the professional development and site visits was estimated at an average of $13,342 per school, $5,132 per teacher, and $80 per student, based on cost data for cohort I. Further, although the courseware was available for free to all Kentucky public schools, the $25 per student registration/support fee assessed by Kentucky Virtual Schools was waived for participating schools. Treatment teachers were not paid for data collection during the intervention year; they did, however, receive a stipend for the time they spent in professional development sessions.[15] Control schools were offered the intervention after their role in the study was complete. Control teachers in year 1 received dinner and reimbursement for local travel for attending an evening orientation meeting, where they were given an overview of their participant responsibilities. Written materials were distributed at these events and mailed to schools that did not attend. Because of the fairly high cost and low participation in the control school orientation meetings, year 2 control schools were informed of their responsibilities by mail.

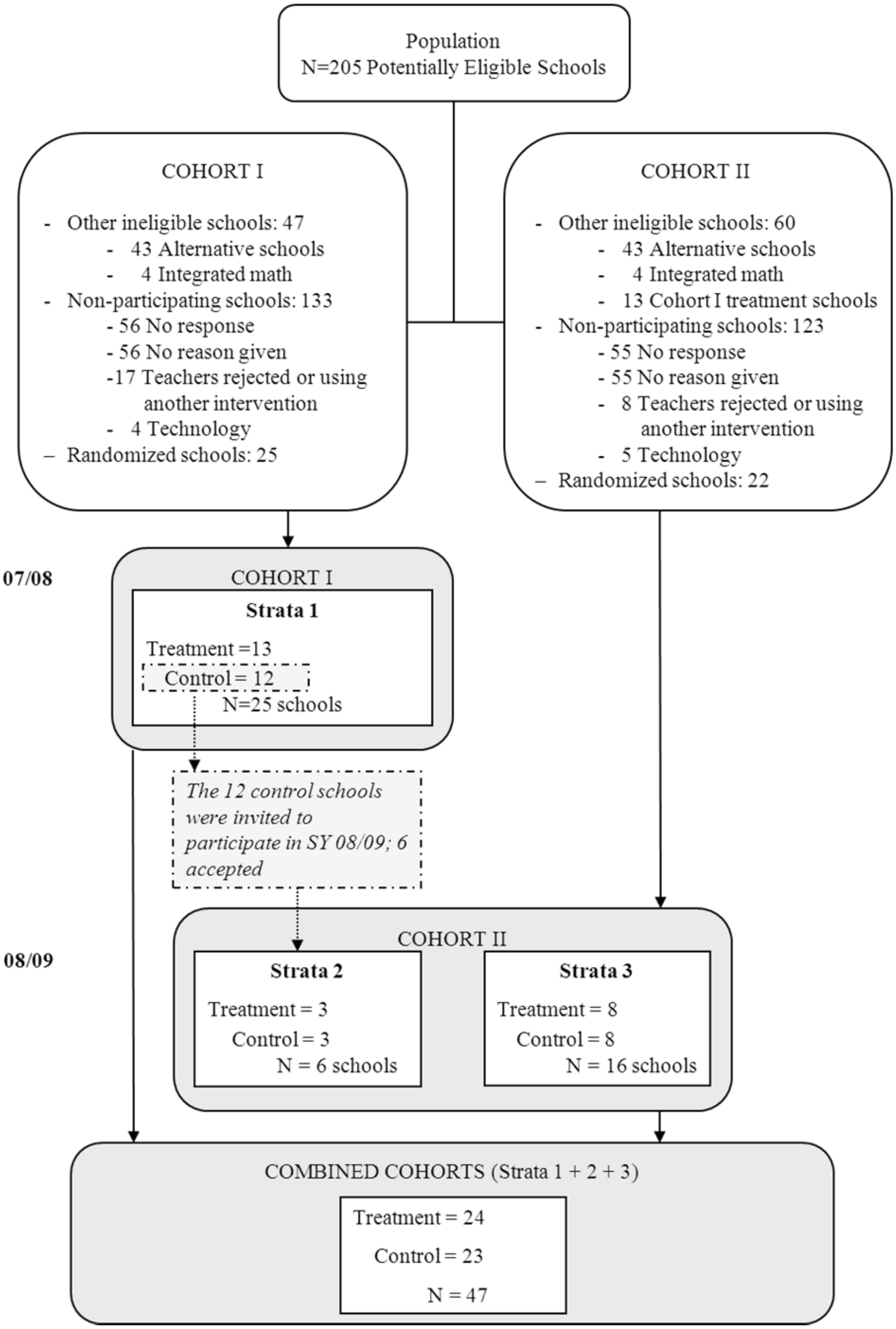
Recruiting challenges

Recruiting schools in rural areas is challenging. Many rural districts have only one high school, which counteracts any advantage to recruiting at the district level. Among the initial list of 205 potentially eligible schools in Cohort I, there were 43 alternative schools and 4 schools that used an integrated math curriculum with no algebra I course in grade 9 that were not eligible to participate (figure 2.1). One hundred and thirty-three schools declined to participate due to the following reasons: 56 schools did not respond to the invitation, 56 schools declined to participate but gave no reason, 17 schools cited teacher objections or had already agreed to participate in an RCT for another algebra I intervention, and 4 schools were concerned that their technology resources could not support the program.[16] The remaining 25 schools were selected for random assignment in cohort I.

[15] A maximum stipend of $1,000 was paid to each teacher participating in professional development. The actual value of the stipend each teacher received was prorated based on attendance at the professional development sessions. Stipends were included in the cost of professional development.

[16] As recruitment for cohort I was being completed, a private company launched a similar RCT—focused on math for grades 8 through 10 and testing Internet-based courseware for use in a classroom setting—in the same geographic area. Note that the counts of reasons schools did not participate are likely understated because of missing data (schools that did not respond or give a reason for declining to participate).

Figure 2.1. Structure of the school sample from random assignment



The same 47 schools that were ineligible to participate in cohort I were also ineligible in cohort II. The 13 treatment schools from cohort I were automatically ineligible to participate in cohort II to avoid treatment-to-control spillover across cohorts and maturation effects from schools participating in the treatment for a second year. Treatment teachers from cohort I received the professional development and had access to the courseware for a year. If they were assigned to the control group in cohort II, the effects of the professional development from participation in cohort I could have spilled over into cohort II. If, on the other hand, the teachers were assigned to a treatment status again in cohort II, they would have had more experience with the intervention, which could confound the treatment effect. However, the control schools in year 1 did not have access to the professional development or the courseware. To increase sample size, all 12 control schools in cohort I were asked to enter cohort II in 2008/09. The actual number of schools eligible to participate in cohort II was 145. Of the 123 eligible schools that did not participate, 55 did not respond to the invitation and 55 schools declined but gave no reason. Three schools cited teachers declining to participate, five cited inadequate technology resources, and five were using another algebra I intervention. The remaining 22, consisting of 6 rerandomized control schools from cohort I and 16 new schools, were selected for random assignment in cohort II. The combined sample with cohort I and cohort II included 47 schools (24 treatment schools and 23 control schools).

Application procedures

Principals of schools wishing to participate in the RCT filled out an application affirming their support for the study. The technology coordinator cosigned the form, verifying that the school had adequate technology and supports to accommodate the technology-based curriculum. These requirements were determined and verified by Kentucky Virtual Schools. The district superintendent also signed the application. In addition to the principal's application form, an application was required from each algebra I teacher within the school. Since this was a whole-school intervention, all algebra I teachers had to agree to participate. Applicants who met these criteria completed a memorandum of understanding, agreeing to furnish all data necessary for analysis, permit one school visit for classroom observations, and posttesting of students' knowledge of algebra I in each year of the study. (See appendix B for more information on this test.)

Random assignment

Randomization occurred at the school level. All algebra I teachers and students in a school were assigned to the school's treatment group, and all algebra I teachers in a treatment school were asked to use the intervention. School-level randomization was chosen, as opposed to within-school randomization, because it minimizes the potential for spillover effects from the treatment to control classrooms which could lead to underestimating the impacts of the hybrid program.

For random assignment, a random number generator for continuous uniform distributions was used to assign a random number to each school and rank the schools according to that number. The first half of the list was assigned to the treatment group, the second to the control group. Cohort I had an odd number (25) of schools to randomize (13 schools were distributed to the treatment group and 12 were distributed to the control

group). For cohort II, the volunteer schools included 16 new schools and 6 control schools from cohort I. The two groups were randomized separately to give each group a 50 percent chance of assignment to the intervention. The combined recruiting sample consisted of three strata from the random assignment process: 25 schools initially assigned to cohort I, 6 schools from the cohort I control group that were rerandomized into cohort II, and 16 new schools randomized into cohort II (see figure 2.1). The combined recruiting sample included 47 schools: 41 unique schools and 6 represented in both cohorts. The treatment group consisted of 24 schools, 63 teachers, and 3,395 students. The control group consisted of 23 schools, 60 teachers, and 3,513 students.

Eligible students

Once the schools were randomized, the grade 9 students eligible for the study sample were identified based on enrollment on September 1 in a course culminating in algebra I credit during that school year. The sample was restricted to students present at the start of the study year, because this group represents the students in each school who were initially randomized into the sample. The restriction ensured that movement of students after the start of the intervention did not disturb the composition of the randomly assigned groups.

The intent-to-treat sample consisted of all grade 9 students in treatment or control schools enrolled on September 1 in a course culminating in the completion of algebra I credit. According to Kentucky guidelines, students are eligible to take algebra I if they did not complete algebra I in grade 8. Data from the Kentucky Department of Education and, in some cases, schools or districts were used to identify eligible students.

State course codes were used to categorize the math courses and the following guidelines were used to determine which were eligible for inclusion in the sample:

Eligible courses:

- *Algebra I (state course code 270304).* This is either a one-year course or a “block-schedule” one-semester course that fulfills the algebra I requirement. Both formats are included in the sample because this study used a whole-school research design, meaning that treatment students in both types of courses were selected to receive the intervention.
- *Algebra I – part B (state course code 270303).* This is the second course in a two-year sequence that fulfills the algebra I requirement.
- *Algebra I honors (state course code 270305).* This course fulfills the algebra I requirement and provides some extensions.
- *Accelerated algebra I (state course code 270306).* This course fulfills the algebra I requirement and provides some extensions and acceleration.

Ineligible courses:

- *Algebra I – part A (state course code 270302).* This is the first course in a two-year sequence. It counts toward elective credit but does not fulfill the algebra I requirement unless Part B is also completed.
- *Algebra I lab (state course code 270308).* This is an elective course that does not count toward algebra I credit. Some eligible state course codes are assigned courses with the name "algebra I lab." It was assumed that any course with a 270308 state course code or "lab" in its name is an algebra I lab course.
- Courses at individual schools with an eligible state course code but that do not fulfill the algebra I requirement or require algebra I as a prerequisite for enrollment (information provided directly by the school).These include KBHC algebra I, VSA algebra I, MA intermediate algebra/introduction to geometry, and algebra I plus.
- Courses taught by an alternative school teacher at a location off-site from the study school.

If a student was concurrently enrolled in an eligible and an ineligible course at the beginning of the school year (for example, algebra I and algebra I lab), the student was included in the sample with the enrollment record from the eligible course. Some students were concurrently enrolled in different types of eligible courses (regular and honors). These students were included in the sample as enrolled in an eligible course, but the course level was categorized as missing because which course level the student actually enrolled in was unknown.

The algebra I courses are offered in three primary formats: full-year (two semesters); block schedules, in which the course is completed in one semester; and two-year, in which half the course is completed in each of two different school years (four semesters; table 2.2). In typical full-year courses (the most common format), students are assigned to a single teacher and course leading to one algebra credit. A full-year course typically runs fall to spring, but it can also run spring to fall, crossing school years. All full-year courses that ended during the intervention year, as well as two-year and single-semester courses that culminated in the completion of algebra I credit during the intervention year, were included in the intent-to-treat sample.[17] Even though the exposure time differed by course format, the curriculum for all of the courses is based on the Kentucky state standards for Algebra I and successful completion of any of these courses fulfills the state graduation requirement for Algebra I. The intervention was designed to be implemented the same way in each course format, and all algebra I teachers in treatment schools were asked to

[17] Students who enrolled in a two-semester or two-year course that started in grade 8 received the intervention for only half the duration of the course. These students were included in the sample because the study used a whole-school design. Systematically removing some courses from the sample could have led to systematic differences between students included in the sample and those omitted, biasing the results of the study. Sensitivity analyses show that there is no statistically significant difference in the impact estimates when students enrolled on September 1 in a part-year course leading to the completion of algebra I credit during the intervention period are excluded from the analysis sample (see appendix H).

use the intervention, regardless of the format.11.28 percent of students in the full sample were enrolled in block courses and 88.79 percent of students were enrolled in full-year courses over two semesters or two years (86.39 percent in full-year courses in grade 9, 0.59 percent in full-year courses spanning grades 8 and 9, and 1.81 percent in two-year courses spanning grades 8 and9) (table 2.2). Compared to students in block courses, students in full-year courses had lower values on the student-level deviation from the school-level average math pretests (KCCT and EXPLORE), were less likely to be enrolled in honors courses, and were more likely to be underserved minorities (Black or Hispanic), males, and recipients of free or reduced-price lunch (table 2.3). Each difference was statistically significant at the 95 percent confidence level, using a two-tailed test.

Table 2.2. Characteristics of algebra I courses for students in grade 9 during the intervention year

Type of algebra I class	**Grade**			**Years**			**Included in the study**			**Percent of all students**
	8	9	10	1/2	1	2	Yes	No	**Percent of sample (N=6,908)**	**enrolled in a study school (N=8,586)**
Full-year		X			X		X		86.39	80.26
Full-year		X	X		X			X	n/a	0.02
Full-year	X	X			X		X		0.59	0.73
Block schedule		X		X			X		11.28	13.23
Two years	X	X				X	X		1.81	4.05
Two years		X	X			X		X	n/a	1.70

Source: Authors' compilation based on Kentucky Department of Education data for enrollment (2007/08 and 2008/09).

Table 2.3. Selected descriptive statistics for student-level covariates, by type of algebra I course format (percent, unless otherwise noted)

	Type of course			
Covariate	**Full-year**	**Block**	**Difference**	***p***
Underserved minority	6.51	2.84	3.67*	<0.01
Male	52.10	48.38	3.72*	0.04
Age (mean)	15.45	15.44	0.01	0.77
Recipient of free or reduced-price lunch	63.87	44.66	19.21*	< 0.01
Enrolled in Individualized Education Plan	10.85	9.15	1.70	0.15
Course level: honors	8.89	12.27	– 3.38*	< 0.01
Average student-level deviation from the school-level average on the KCCT math pretest (mean)	–0.58	4.46	– 5.04*	< 0.01
Average student-level deviation from the school-level average on the EXPLORE math pretest (mean)	–0.09	0.68	– 0.77*	< 0.01
Number of students	6,093	815		

* The difference is statistically significant at the 95 percent confidence level using a two-tailed test.
Note: The *p*-values are based on a *t*-test of the difference between students in full-year and block format courses.
Source: Authors' calculations based on Kentucky Department of Education data for demographic characteristics (2007/08 and 2008/09), Kentucky Department of Education Commonwealth Accountability testing system results for KCCT (2006/07 and 2007/08) and EXPLORE (2006/07 and 2007/08), and data provided directly by schools (2007/08 and 2008/09).

Attrition and noncompliance

School attrition and noncompliance

The intent-to-treat sample included all schools assigned to use the hybrid program, regardless of whether the intervention was used as intended. Because the sample was randomized at the school level, the consequences of school attrition were carefully considered. At the school level, researchers planned for attrition of as much as 20 percent. No schools were lost from the analysis sample, however, because all outcome data for the analyses were obtained from the Kentucky Department of Education. On the other hand, four treatment schools were noncompliant (did not participate in the intervention). Reasons for noncompliance included schools that never hired a permanent teacher for algebra I, that participated in a competing study with another type of algebra I intervention immediately after randomization, or that had teachers who refused to participate because they reported being given the application to sign without sufficient information on the agreement (the number of schools in each condition is too few to report due to potential disclosure risk). The intent-to-treat sample included all schools in the study sample, regardless of their compliance status.

Teacher attrition and noncompliance

The instructional specialists responsible for professional development monitored and helped ensure continuing teacher participation. Because the intervention involved scheduled and on-demand support from an instructional specialist (through regularly scheduled sessions with their learning community and site visits), significant attrition was not anticipated from teachers who remained in their schools.

Attrition was also minimized because of the stipends provided to treatment teachers to compensate for their time spent in the professional development sessions. In implementation, there were no losses to the analysis sample due to teachers opting out of testing because the student assessment data were furnished by the Kentucky Department of Education. However, four teachers in the math achievement analysis sample and 5 teachers in the grade 10 math course enrollment sample did not have any students with outcome data. The reasons for these missing data are unknown. In addition, some teachers in the treatment group did not use the intervention (see chapter 4). All treatment teachers were included in the analysis of the fidelity of implementation, even if they did not have any students with outcome data or they did not use the intervention.

Student attrition

Only students enrolled in an algebra I course on September 1 of the intervention year were included in the analysis of student performance. Students who dropped the course continued to be considered part of the study and were handled as if they had remained in their treatment condition. Students who added the course after September 1 were not part of the randomized sample and were excluded from the main confirmatory analysis.

Because the PLAN is given in the first part of September of grade 10, students who did not advance to grade 10 or attend school on the testing day (N=282 control students and 299 treatment students), or who left Kentucky's public school system (N=56 control students and 70 treatment students), were lost from the sample. There were also 158

control students and 179 treatment students with missing PLAN math achievement test data or enrollment data in year 2 for unknown reasons and excluded from the ITT impact analyses of the PLAN outcome. In the overall sample, 86 percent of the randomized group of students in the control group and 84 percent in the treatment group were retained. Attrition was five percentage points lower in cohort I than in cohort II (cohort I = 13 percent; cohort II = 18 percent; $p < .01$).

For the grade 10 math course enrollment outcome, students who left Kentucky's public school system were lost from the sample (N=41 control students and 47 treatment students). There were also 205 control students and 206 treatment students with no enrollment data in year 2 who were missing outcome data for unknown reasons and excluded from the impact analyses of the grade 10 math course enrollment outcome. In the overall sample, 93 percent of students in both the treatment and control groups had data for this outcome. For the grade 10 math course enrollment, attrition was three percentage points lower in cohort I than cohort II (cohort I = 6 percent; cohort II = 9 percent; $p < .01$).

Identifying crossovers

Study participants who changed treatment status during the study were identified as crossovers. In each cohort, there were no crossovers at the teacher or school level. The number of crossovers was minimized because the unit of randomization was the school, rather than the classroom. A student was categorized as a crossover if he/she transferred to a different study school and school treatment condition, or to a nonstudy school. To identify student crossovers, the schools' student enrollment records from September 1 to May 1 were compared. In the combined cohorts, 557 grade 9 students in eligible courses (7.95 percent) transferred to a school that resulted in a change of treatment or control status (table 2.4). The changes were handled as follows:

- The largest group of students who switched status transferred from a study school to a nonstudy school: 257 students switched from a control school, and 192 students switched from a treatment school. Typically, students who transfer from a control school to a nonstudy school are considered a natural part of the counterfactual. However, the hybrid algebra courseware was delivered in schools outside the study, so transferring to a nonstudy school could entail crossover.[18] All students who transferred from a study school to a nonstudy school were included in the sample with the treatment status from the initial enrollment record.
- Ninety-eight students started at a nonstudy school and transferred to a control school (n = 47) or a treatment school (n = 51). These students were not included in the analysis sample because on September 1 they were enrolled in a nonstudy school.

[18] During the 2007/08 and 2008/09 school years, 3,717 students were enrolled in Kentucky Virtual Schools' hybrid courses across all subjects and grade levels. (http://education.ky.gov/kyvsassets/kvhs%20enrollment%20and%20trend%20data.xls). No data were available on the number of students who switched from a control school to a nonstudy school that used the hybrid algebra courseware.

- Only 10 students started at one study school and transferred to another study school with a different treatment condition. All these students were included in the sample with their initially assigned treatment condition.

Table 2.4. Students in the full sample who switched treatment or control status during the intervention school year, by type of transfer

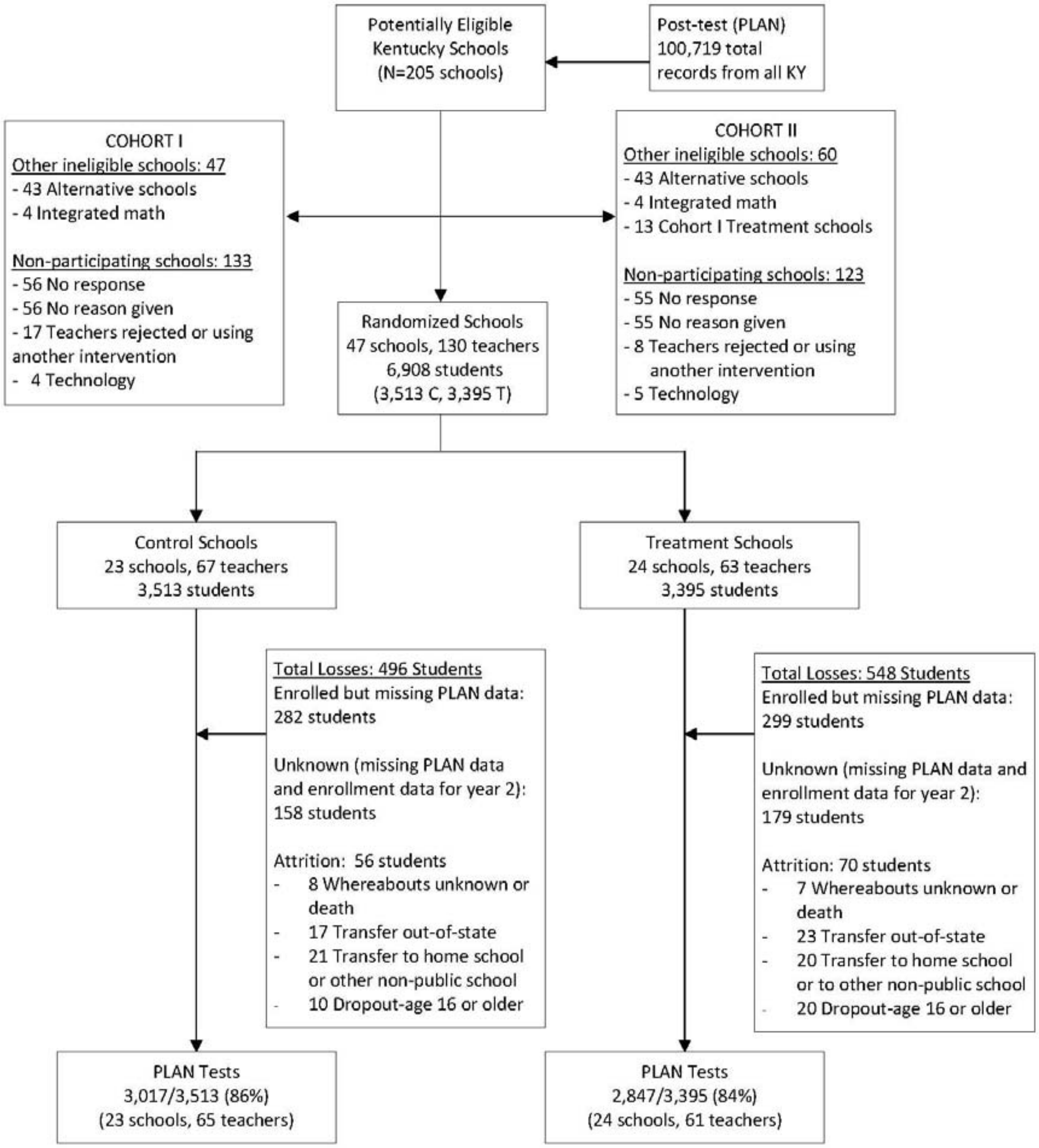
Type of transfer	*n*	*Percent*
Control to nonstudy school	257	3.67
Treatment to nonstudy school	192	2.74
Nonstudy to control school	47	0.67
Nonstudy to treatment school	51	0.73
Control to treatment school or treatment to control school	10	0.14
Total change in status	557	7.95

Note: N=7,006 grade 9 students who were enrolled in an eligible course and had an enrollment record at a study school at any time during the intervention year.
Source: Authors' analysis of Kentucky Department of Education data for student enrollment (2007/08 and 2008/09).

Study sample

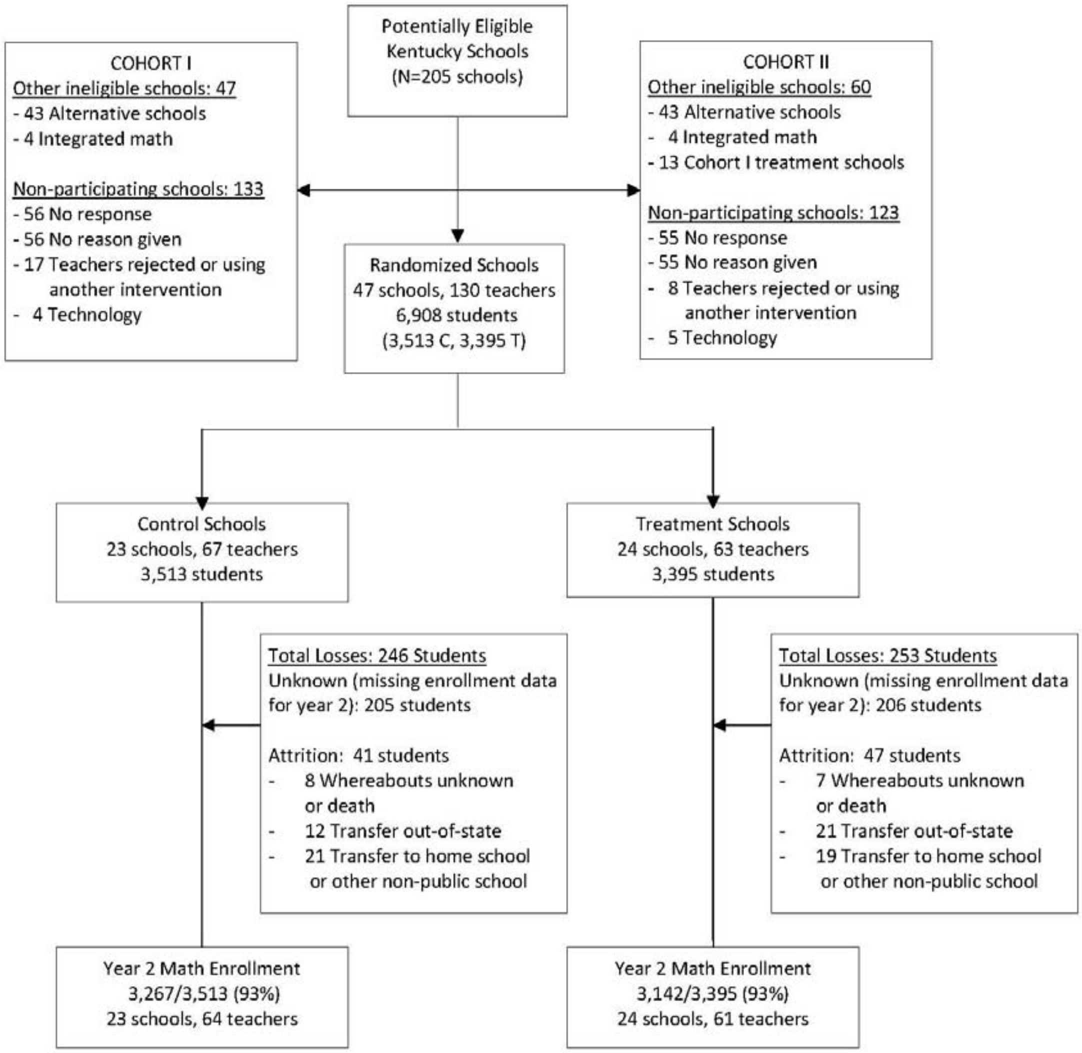
The assignment of study participants and participant losses is summarized in flowcharts adapted from the Consolidated Standards on Reporting Trials (CONSORT) statement (Altman et al. 2001). Figures 2.2 and 2.3 display the sample detail for the PLAN math achievement sample and the sample detail for the grade 10 math course enrollment sample. There were 6,908 grade 9 algebra I students in participating schools (the baseline sample). The response rates for the PLAN were similar for the control group (86 percent) and the treatment group (84 percent). Four teachers were lost from the analysis sample; all of these teachers had three or fewer students and these students were missing data for one of the reasons listed in the "total losses" box. In both the treatment and control groups, 93 percent of students in the sample had outcome data for grade 10 math course enrollment. Five teachers were lost from the analysis sample. It is unknown why these teachers were missing outcome data. See appendix C for separate figures of sample detail for cohorts I and II for each outcome.

Figure 2.2. Sample detail for the grade 10 PLAN assessment of pre-algebra/algebra skills



Source: Authors' calculations based on Kentucky Department of Education data for student enrollment (2007/08, 2008/09, and 2009/10) and Kentucky Department of Education Commonwealth Accountability testing system results for the PLAN (2008/09 and 2009/10). Adapted from the Consolidated Standards on Reporting Trials (CONSORT) statement (Altman et al. 2001).

Figure 2.3. Sample detail for grade 10 math course enrollment



Source: Authors' calculations based on Kentucky Department of Education data for student enrollment (2007/08, 2008/09, and 2009/10). Adapted from the Consolidated Standards on Reporting Trials (CONSORT) statement (Altman et al. 2001).

Baseline characteristics of the study sample

With a couple of exceptions, the study sample is fairly typical of ninth grade students in the Appalachia region. Relatively low percentages are from underserved minority groups; roughly half are males; the mean age is about 15.5 years; and about 10 percent have individual education plans (Table 2.5). Notably, a majority of the students in the study sample are eligible for free or reduced price meals and low percentages are enrolled in honors courses. Over half of the schools are located in rural areas. Since the recruitment of schools was limited to those in which a maximum of 60 percent of students were proficient in math, the school-level average pretest scores for the sample

are lower than the state average on the EXPLORE math (state average=14.20) and the KCCT math (state average=839.58).

To determine if the differences between the treatment and control groups were statistically significant, ordinary least squares regression was conducted, where the dependent variable was a dichotomous indicator for treatment status. The standard errors for school-level characteristics were adjusted to account for clustering of residuals at the school level and only one statistically significant difference between the treatment and control groups was found in the randomized sample: students in the treatment group were more likely to be male than students in the control group.[19]

Table 2.5. Baseline characteristics in the treatment and control groups for the full study sample

	Study sample			
	Treatment	**Control**	**Difference**	***p***
Student covariates				
Underserved minority (percent)	4.19	7.92	–3.73	0.05
Male (percent)	53.08	50.37	2.71*	0.02
Age (mean)	15.48	15.40	0.08	0.59
Recipient of free or reduced-price lunch (percent)	63.32	60.00	3.32	0.09
Enrolled in Individualized Education Plan (percent)	10.67	10.63	0.04	0.92
Course level: honors (percent)	6.37	12.11	–5.74	0.33
Student-level deviation from the school-level mean on the KCCT math pretest	0.00	0.00	0.00	0.54
Student-level deviation from the school-level mean on the EXPLORE math pretest deviation	0.00	0.00	0.00	0.10
Sample size	3,395	3,513		
School covariates				
School-level KCCT math pretest score (mean)	836.85	837.42	–0.57	0.82
School-level EXPLORE math pretest score (mean)	13.86	13.82	0.04	0.95
Rural (percent)	53.93	68.60	–14.67	0.50
Strata 1 (cohort I schools) (percent)	48.28	57.50	–9.22	0.72
Strata 2 (rerandomized cohort I schools) (percent)	6.75	12.75	–6.01	0.50
Strata 3 (new cohort II schools) (percent)	44.98	29.75	15.23	0.77
Sample size	24	23		

* The difference is statistically significant at the 95 percent confidence level using a two-tailed test.
Source: Authors' calculations based on Kentucky Department of Education data for demographic characteristics (2007/08 and 2008/09), Kentucky Department of Education Commonwealth Accountability testing system results for KCCT (2006/07 and 2007/08) and EXPLORE (2006/07 and 2007/08), and data provided directly by schools (2007/08 and 2008/09).

Because randomization occurred at the school level, the gender difference between the treatment and control group was attributed to the characteristics of a few large schools. The control group had four large schools ($n > 200$) with a gender imbalance (47.1 percent male and 51.5 percent for the remaining control schools). Researchers

[19] As a sensitivity test, pooled cohort statistical test of baseline characteristics for the sample of 41 schools was rerun. This sample excluded the six rerandomized schools in strata 2. There were no statistically significant differences in any of the baseline characteristics for this sample.

controlled for this covariate in the impact analysis to remove any bias caused by these differences, to reduce unexplained variation in outcomes, and to improve the precision of impact estimates.

We also tested for differences in baseline characteristics by treatment status for students who were missing outcome data (results not shown here). Students in the treatment group were less likely to be underserved minorities (Black or Hispanic) among those missing PLAN data (difference=-2.71, p=0.03) and those missing year 2 math course enrollment data (difference=-3.28, p=0.01). Students in the treatment group were also less likely to be enrolled in an honors algebra I class among those missing PLAN data (difference=-3.28, p=0.01) and those missing year 2 math course enrollment data (difference=-5.26, p=0.03). To minimize potential bias, statistical adjustments were made to the impact estimates using ordinary least squares regression adjustment for these baseline covariates that differ by treatment status for students missing outcome data.

3. Data and Methods

This chapter begins with a discussion of the data used in the study, including the outcome measures and covariates used in the impact analysis, as well as the observational and survey data used to examine the implementation of the intervention. It then describes the methods used to estimate impacts, including the procedures for handling missing data, conducting sensitivity analyses, adjusting for multiple comparisons, and weighting.

Data

Table 3.1 summarizes the data collection process for each type of measure used in the analysis. (Additional data collected but not used are detailed in appendix B.[20])

[20] The Kentucky Department of Education provided PLAN composite scores in four subject areas (English, math, reading, and science). Only the algebra/pre-algebra strand of the math assessment was used for this evaluation, due to the tenuous connection between the intervention and the other subject areas. An end-of-course assessment was conducted at the end of the intervention year but was excluded from the analyses because of high nonresponse rates (29 percent of students) and differential attrition (4 percentage points higher in the treatment group than in the control group). Although these attrition rates met Institute of Education Sciences evidence standards for outcome nonresponse, the PLAN data, with an 85 percent overall response rate and a 2 percentage point gap in attrition between treatment and control groups, presented less of a threat to the internal validity of the impact estimates. The end-of-course assessment was also given in the postintervention year to evaluate teacher effectiveness but, given the high nonresponse rates and differential attrition, was not given to participants in the second year for cohort II and is not included in the analyses. Also, the Kentucky Department of Education provided enrollment and withdrawal data in year 2 for cohorts I and II to assess students' educational persistence, but this analysis was dropped from the study design because of the hypothesized weak connection with the intervention.

Table 3.1. Data collection summary

Type of measure	Instrument	Timeline	Number of cases with data	Description
Math achievement; fall of grade 10	American College Testing (ACT) PLAN: pre-algebra/algebra portion	September 2008 (cohort I), September 2009 (cohort II)	3,185 (cohort I) 2,679 (cohort II) 5,864 (combined)	State administered to all grade 10 students. Response rate 85 percent.
Grade 10 math course enrollment	Math course enrollment records in the post-intervention year	2008/09 (cohort I), 2009/10 (cohort II)	3,437 (cohort I) 2,972 (cohort II) 6,409 (combined)	State enrollment records collected in the spring of the postintervention year. Response rate 93 percent.
Teacher surveys[21]	Hybrid Algebra 1 Teacher Questionnaire for treatment teachers	Spring 2008 (cohort I), Spring 2009 (cohort II)	43 (combined cohorts)	All teachers were asked to complete the mail survey. Response rates: 68 percent of treatment teachers and 70 percent of control teachers
	Algebra 1 Control Teacher Questionnaire for control teachers		47 (combined cohorts)	
Classroom observations	School Observation Measure	2007/08 (cohort I), 2008/09 (cohort II)	165 (combined cohorts)	165 observations by trained site researchers
	Algebra I Quality Assessment		165 (combined cohorts)	165 observations: 80 in treatment schools and 85 in control schools. Response rates: 60 percent for treatment teachers and 66 percent for control teachers

Source: Authors' compilation based on Kentucky Department of Education data for enrollment (2008/09 and 2009/10), Kentucky Department of Education Commonwealth Accountability Testing System results for the PLAN (2007/08 and 2008/09), the Hybrid Algebra 1 Teacher Questionnaire (2007/08 and 2008/09), Algebra 1 Control Teacher Questionnaire (2007/08 and 2008/09), the School Observation Measure (2007/08 and 2008/09), the Algebra I Quality Assessment (2007/08 and 2008/09), and data provided directly by schools (2007/08 and 2008/09).

Outcome measures

Scale score on the pre-algebra/algebra strand of the American College Testing PLAN assessment

The effect of the intervention on student posttreatment math knowledge and skills was assessed with the pre-algebra/algebra portion of the American College Testing (ACT) PLAN. This is a curriculum-based test that covers the skills and knowledge that are commonly taught in the nation's schools and are judged to be important for success in both high school and college, which is administered to all Kentucky students in grade 10. The tests measure what students know and what they are able to do with their knowledge" (ACT 2009). According to the Kentucky Department of Education, the PLAN is closely aligned with the Department's learning goals and curriculum standards

[21] Any inquiries regarding the teacher surveys or classroom observation instruments and methodology used for this project may be directed to: The Center for Research in Educational Policy, The University of Memphis, 325 Browning Hall, Memphis, TN 38152.

for high school math. Of the PLAN's 40 items, 22 cover pre-algebra and algebra topics, and 18 cover geometry. The scale score on the pre-algebra/algebra strand of the PLAN was the dependent variable in the primary confirmatory analysis of student achievement.

The PLAN's pre-algebra/algebra scale score was selected for the confirmatory analysis because it aligns better with the intervention than with the assessment's composite math score that includes geometry. According to the ACT, and based on a national sample, the pre-algebra/algebra strand has an estimated reliability of 0.78 and a standard error of measurement of 1.62, compared with a reliability estimate of 0.81 and a standard error of measurement of 1.98 for the composite math scale score (ACT 2007a). All tests were scored by ACT, Inc. Approximately 6 percent of treatment students and 6 percent of control students received test accommodations that included extended time, cassette tests, reader scripts, raised line drawings, Braille tests, and/or assistive communication devices.

Because the PLAN was administered in the fall semester following the intervention, short-term impacts that did not persist until fall will not be captured in the analysis.[22] However, there should not be any bias in the estimate of impacts that persisted until fall of tenth grade.

Grade 10 math course enrollment

Grade 10 math course enrollment is a dichotomous indicator for whether a student's math course enrollment in grade 10 was above algebra I, with students enrolled in a more advanced math course receiving a value of 1 on this measure and other students receiving a value of 0. Information to construct this measure was obtained from the Kentucky Department of Education through separate data files containing information school enrollment, math course enrollment, and withdrawal for students in the study sample. Department of Education records were incomplete for four of the 47 schools in the study sample. However, these schools provided data directly to the study team. Below is a detailed description of how this variable was constructed, with the sample sizes and the percentages of all students in each cohort with the corresponding values:

Grade 10 math course enrollment = 1

> Students enrolled in a math course(s) above algebra I in any semester of the postintervention school year (state course code of 270310 or above) and were not enrolled in algebra I or pre-algebra. Courses above algebra I include algebra II, trigonometry, geometry, calculus, and "other mathematical topics" (probability and statistics, mathematics for business and industry, technical mathematics; cohort I: n = 2,865, 78 percent; cohort II: n = 2,515, 77 percent).

Grade 10 math course enrollment = 0

> Students enrolled in an algebra I or pre-algebra course (state course code of 270308 or below). This value is assigned to all students enrolled in algebra I or

[22] A meta-analysis of the effects of summer vacation on achievement test scores (Cooper et al. 1996) found that the scores declined about one-tenth of a standard deviation between the end of one school year and the start of the next, and the effect of summer break was more detrimental for math than for reading.

pre-algebra in any semester of the postintervention school year, regardless of whether the student is also enrolled in a course above algebra I (cohort I: *n* = 492, 13 percent; cohort II: *n* = 369, 11 percent).

Students enrolled in a Kentucky public school but did not enroll in any math course in the postintervention school year (cohort I: *n* = 73, 2 percent; cohort II: *n* = 64, 2 percent).

- Students who dropped out of school at the beginning of the postintervention school year (cohort I: *n* = 7, 0.2 percent; cohort II: *n* = 24, 0.7 percent).

Grade 10 math course enrollment = missing

- Students with no math course enrollment records but with a withdrawal record other than dropout, indicating that they were no longer enrolled in a Kentucky public school. Such students may have transferred to a private school, out-of-state school, or home school or had withdrawal codes indicating "moved, whereabouts unknown," "withdrawn due to health reasons," or "withdrawn due to death" (cohort I: *n* = 23, 1 percent; cohort II: *n* = 67, 2 percent).
- Students with an unknown enrollment status in the postintervention school year (no records in the school enrollment, math course enrollment, or withdrawal files; cohort I: *n* = 199, 5 percent; cohort II: *n* = 210, 6 percent).

Covariates

The analysis included pretest achievement measures and other covariates to reduce the confidence intervals surrounding the point estimates of impacts. Some of them also were used to construct subgroups for the exploratory analysis.

Pretests

Two tests of math knowledge and skills given to all Kentucky students in grade 8 were used as pretest measures in the confirmatory and exploratory analyses: the KCCT (Kentucky Department of Education 2009) and the American College Testing EXPLORE (ACT 2007b). Two covariates were included for each pretest: (1) the average pretest at the school level and (2) the student-level deviation from the school-level average pretest. The first covariate reduces the school-level variance; the second reduces within-school variance. These variables were created by calculating the average scale score for students in the sample at each school and subtracting the students' scale scores from the school-level average.

Kentucky Core Content Test (KCCT) for grade 8 (math)

The KCCT, which assesses students in content areas specific to their grade level, is a major component of the Commonwealth Accountability Testing System. Its test results are used to evaluate the school program in the state accountability system. The results from the reading and math content areas are also used to meet federal testing and reporting requirements of the No Child Left Behind Act of 2001. The criterion-referenced

test items measure Kentucky Core Content for Assessment (v4.1), a subset of the Program of Studies.[23] The KCCT math test has an estimated reliability of 0.89 and a standard error of measurement of 3.29, based on the overall population of students in grade 8 who took the 2007/08 KCCT assessments (Kentucky Department of Education 2009). The grade 8 math scale score on the KCCT was the first source for the pretest covariates.

American College Testing EXPLORE: College Readiness Test for grade 8 (math)

The grade 8 math scale score on the EXPLORE was the second source for the pretest covariates. EXPLORE is a curriculum-based assessment program that helps students in grade 8 understand their academic development in college preparation, make the most of their opportunities in high school and beyond, and guide them as they start thinking about future education and career planning. EXPLORE assesses academic progress, provides an early indicator of college readiness, helps students understand and begin to explore the wide range of career options open to them, and assists them in developing a high school coursework plan that prepares them to achieve their post–high school goals. All students enrolled in a Kentucky public school in grade 8 (except those enrolled in the Alternate Assessment program)[24] complete the EXPLORE assessment. EXPLORE includes four 30-minute multiple-choice tests—English, math, reading, and science—and collects information on students' interests, needs, plans, and selected background characteristics. The math portion consists of 30 items testing the following areas: pre-algebra (10 items), elementary algebra (9 items), geometry (7 items), and statistics/probability (4 items). According to ACT, the EXPLORE math test has an estimated reliability of 0.82 and a standard error of measurement of 1.71, based on a national sample of grade 8 students (ACT 2007b).

Other student-level covariates

The models for both student outcomes included the following additional student-level covariates:

- Free or reduced-priced lunch status: equal to 1 if the student was a recipient and 0 if the student was a nonrecipient.
- Gender: equal to 1 if the student was male and 0 if the student was female.
- Underserved minority status: equal to 1 if the student was Black or Hispanic and 0 if the student belonged to another racial/ethnic group.
- Individualized Education Plan status: equal to 1 if the student had an Individualized Education Plan and 0 if the student did not.

---

[23] The Program of Studies is the minimum required content standard students shall be taught to meet the high school graduation requirements (http://www.education.ky.gov/kde/instructional+resources/curriculum+documents+and+resources/program+of+studies/).

[24] The Kentucky Alternate Assessment Program is for students with moderate to significant disabilities and for whom traditional assessments would be an inappropriate measure of progress. Less than 1 percent of the total student population in the state participates in the program (http://www.education.ky.gov/KDE/Administrative+Resources/Testing+and+Reporting+/District+Support/Kentucky+Alternate+Assessment+Program/).

- Age in years on September 1 of the intervention year. This was a continuous variable derived from the student's date of birth.
- Course level: 1 if the student was enrolled in an honors-level algebra I course and 0 if the student was enrolled in a regular algebra I course (all students were enrolled in a regular or honors-level algebra I course).

School-level covariates

The models for both student outcomes included additional school-level covariates, defined as follows:

- Rural status: equal to 1 for rural schools and 0 for nonrural schools.
- Strata: a pair of dichotomous variables denoting the school clusters for random assignment. Strata 2 was equal to 1 for the six schools from the cohort I control group that were rerandomized into cohort II in 2008/09 and 0 otherwise. Strata 3 was equal to 1 for the 16 new schools that participated in the study in 2008/09 and 0 otherwise. For the exploratory analysis of whether there were statistically significant differences between the impacts of the intervention by cohort, a single dichotomous variable for cohort I (0 = cohort II, 1 = cohort I) was substituted for the strata variables.

Collection of data for the impact evaluation

The Kentucky Department of Education provided data for the PLAN posttest for students attending all Kentucky public schools. The PLAN was administered to grade 10 students in the first half of September of their fall semester, according to state guidelines. The department also provided the data for constructing the outcome variable for grade 10 math course enrollment. The school enrollment, math course enrollment, and withdrawal records were collected during the spring semester of the postintervention year. The math course enrollment file included courses for which the student was enrolled in September for the fall semester and in January for the spring semester. In practice, the first grade 10 math enrollment record for the sample was in the fall semester for 92 percent of students and in the spring semester for 8 percent of students.[25]

All students in grade 8 take EXPLORE during their fall semester and the KCCT during their spring semester, according to state guidelines. The Kentucky Department of Education provided data for both KCCT and EXPLORE. For cohort I, data for both pretests were obtained one year before the intervention, meaning that any student who repeated grade 9 in the intervention implementation year was missing pretest data. For cohort II, there were data for both pretests for one and two years before the intervention. Of the sample, 89.6 percent had records for both pretests (table 3.2). Approximately 6.5

[25] There was no statistically significant difference ($p = .21$) for the timing of the first grade 10 math enrollment record between the treatment and control groups. In cohort I, the first grade 10 math enrollment record was in the fall semester for 91.6 percent of students in the control group and 92.8 percent of students in the treatment group. Comparable data on the start date of classes were not available for cohort II.

percent of the treatment group and 5.4 percent of the control group had no records for either pretest. Some of these students may not have been enrolled in a Kentucky public school in grade 8; others may have been absent from school on testing days. For cohort I only, students may be missing pretest data because they were not enrolled in grade 8 during the year before the intervention. How many students were missing grade 8 test scores for each reason was unknown. Students with missing pretest records were included in the impact analyses using the dummy variable adjustment method for missing data (Cohen and Cohen 1983).

Table 3.2. Pretest records for the baseline sample, by treatment status

	All	Control	Treatment
Both pretests			
Number	6,187	3,134	3,053
Percent	89.6	89.2	89.9
KCCT pretest only			
Number	260	160	100
Percent	3.8	4.6	2.9
EXPLORE pretest only			
Number	51	28	23
Percent	0.7	0.8	0.7
Missing pretest			
Number	410	191	219
Percent	5.9	5.4	6.5
Total	6,908	3,513	3,395

Source: Authors' calculations based on Kentucky Department of Education Commonwealth Accountability Testing System results for KCCT (2006/07 and 2007/08) and EXPLORE (2006/07 and 2007/08).

Student covariates for gender, race, Individualized Education Plan status, free or reduced-price lunch status, and age were collected from four different sources: (1) the Kentucky Department of Education, (2) KCCT, (3) school data, and (4) EXPLORE. Table 3.3 presents the number of records from each data source.

The Kentucky Department of Education was the primary data source because it is the official state record and was provided with the enrollment records. Students who had pretest data but were missing any of the descriptive covariates were included in the sample using the dummy variable adjustment method for missing data (Cohen and Cohen 1983).

For the gender, race, and age variables, students were assigned values based on the first data source with nonmissing values. For the free or reduced-price lunch and Individualized Education Plan variables, students were assigned the value 1 if any of the data sources categorized them as a recipient of free or reduced-price lunch or enrolled in an Individualized Education Plan. (See appendix E for a description of problems encountered with these variables during data cleaning.)

Table 3.3. Data sources and records for student covariates for the baseline sample

Student covariate	Data source	Number	Percent
Gender	KDE	6,785	98.2
	KCCT or school	115	1.7
	Missing	8	0.1
	Total	6,908	100.0
Race	KDE	6,741	97.6
	KCCT	144	2.1
	School or EXPLORE	3	0.0
	Missing	20	0.3
	Total	6,908	100.0
Individualized Education Plan status	KDE	3,313	48.0
	KCCT	3,380	48.9
	School	40	0.6
	Missing	175	2.5
	Total	6,908	100.0
Free or reduced-price lunch status	KDE	3,322	48.1
	KCCT	3,156	45.7
	School	35	0.5
	EXPLORE	232	3.4
	Missing	163	2.4
	Total	6,908	100.0
Age	KDE	6,742	97.6
	KCCT	143	2.1
	School or EXPLORE	3	0.0
	Missing	20	0.3
	Total	6,908	100.0

KDE = Kentucky Department of Education.
Source: Based on Kentucky Department of Education data for enrollment (2007/08 and 2008/09) and demographic characteristics (2007/08 and 2008/09), Kentucky Department of Education Commonwealth Accountability Testing System results for KCCT (2006/07 and 2007/08) and EXPLORE (2006/07 and 2007/08), and data provided directly by schools (2007/08 and 2008/09).

Algebra I course information was also collected at the student level from the Kentucky Department of Education enrollment records. All eligible algebra I courses were categorized as regular or honors level. Algebra I – part B (state course code 270303) and algebra I (state course code 270304) were coded as regular courses. Algebra I honors (state course code 270305) and accelerated algebra I (state course code 270306) were categorized as honors courses because they provide extensions and or/acceleration. Discussions with the study schools indicated little difference between “honors” courses and “accelerated courses.” The difference seems attributable to the labeling preferences of the individual school.

The indicator for rural schools was based on the school locale variable in the publicly available Common Core of Data from the U.S. Department of Education's National Center for Education Statistics (http://nces.ed.gov/ccd/). The following values for the school locale variable were categorized as rural: rural: fringe (locale code 41), rural: distant (locale code 42), and rural: remote (locale code 42). All other locale codes were categorized as nonrural.

Teacher survey

Teacher surveys and classroom observations were used to describe differences in the treatment and control conditions during the intervention. Teacher surveys were mailed to each treatment and control teacher mid-spring of the intervention year. Teachers were instructed to return the survey to the site researcher during the one-day school visit or in a postage-paid return envelope provided with the survey.

Teacher surveys were used to collect teacher perceptions of their algebra I approach (hybrid or district curriculum). Coburn (2003) suggests that "ownership over the reform must shift so that it is no longer an "external" reform, controlled by a reformer, but rather becomes an "internal" reform with authority for the reform held by districts, schools, and teachers who have the capacity to sustain, spread, and deepen reform principles themselves" (p. 7). Two versions of the survey instrument were used: The Hybrid Algebra 1 Teacher Questionnaire (Hybrid Teacher Questionnaire) and the Algebra 1 Control Teacher Questionnaire (Control Teacher Questionnaire). These instruments were identical except for the terminology referring to the teacher's algebra I approach. The treatment teacher survey uses "hybrid algebra I approach," whereas the control teacher survey uses "my district's algebra I curriculum."

In the first section, teachers rated their level of agreement with 16 statements focusing on three subscales: impact of the algebra I approach on classroom students, impact of the algebra I approach on instruction, and teacher readiness to implement the algebra I approach. Items are rated on a five-point Likert-type scale that ranges from (1) strongly disagree to (5) strongly agree.

On the second section of each questionnaire, teachers indicated the frequency (1 = never; 5 = extensively) with which they used 11 research-based strategies divided into two subscales: teacher use of strategies (3) and student use of strategies (8). The strategies were emphasized in the hybrid algebra I program and were based on National Council of Teachers of Mathematics (NCTM 1989, 1991, 2000), North American Council for Online Learning (NACOL 2007), National Education Technology Standards (ISTE 1998, 2000), and Kentucky algebra I standards (Kentucky Department of Education 2006b). The teacher strategies were use of higher-order questioning, demonstrating algebra concepts, and use of computers to explain algebra. Student strategies were working in groups, using writing, talking, hands-on objects, activities, graphing calculators, computers, and exit slips to explain and learn algebra.

Teachers then were asked to indicate the number of math-related professional development activities (other than hybrid/Spotlight on Algebra I—used for treatment teacher professional development—see chapter 4) they had completed in the last 12 months and the degree to which the professional development changed their algebra I teaching practices (not at all, some, a lot). The survey provided space for teachers to

indicate their perceptions of the greatest benefits, most difficult aspects, areas of needed improvement and their preference for teaching algebra I again with the approach they used to teach it the past year.

The Hybrid Teacher Questionnaire and Control Teacher Questionnaire are an adaptation of the 20-item validated Teacher Technology Questionnaire, designed to assess teacher perceptions of technology integration on five subscales: impact of technology on students, impact of technology on instruction, readiness to integrate technology, technical support, and overall support (Lowther and Ross 2000). The range of reliability indices across all scales was between r=0.38 and r=0.81. The low reliability for some of these scales limits the ability to draw conclusions about the teachers' instructional strategies. See appendix D for a discussion of the reliability calculations for the questionnaires.

Classroom observations

Classrooms were observed during the intervention year using two instruments: (1) the School Observation Measure and (2) the Algebra I Quality Assessment. The School Observation Measure (Ross, Smith, and Alberg 1999) data describe whether there are more teacher-centered strategies (higher-level instructional feedback, teacher as a facilitator, computer for instructional delivery) in control classes and more student-centered strategies (cooperative learning, higher-level questioning, hands-on learning, student discussion, technology as a learning tool, higher level of academic focus, and higher level of student attention/interest) in treatment classes. Algebra I Quality Assessment (Lowther 2006) data identify the frequency and quality of use of the same 11 algebra I strategies on the teacher surveys and the availability of computers for student use in the observed setting. Each item includes a two-part rating scale. The first is a yes/no indicator of whether an activity was observed and the second is a three-level indicator of quality (low, moderate, high). One Algebra I Quality Assessment was completed as part of the observation in each class.

Trained and unbiased observers conducted the School Observation Measure and Algebra I Quality Assessment observations during direct classroom observations of algebra I lessons (approximately one-hour-long). As many as six School Observation Measure/Algebra I Quality Assessment observations were conducted during one-day visits to each treatment and control school. The goal was to observe all algebra I teachers at least once. However, not all teachers in the sample were able to be observed for reasons including the fact that not all teachers were present on the observation day, school administrators did not provide observers with complete rosters of all algebra I teachers, and four schools dropped out of the study. In the control group, 44 of the 67 control group teachers (65.7 percent) and 38 of the 63 teachers treatment group teachers (60.3 percent) were observed. The conclusions that can be drawn about the teachers' instructional strategies are limited both because not all teachers were observed and because most teachers who were observed had only a one day observation.

In a 1999 reliability study of the School Observation Measure reported by Lewis, Ross, and Alberg, pairs of trained observers selected the identical overall response on the five-category rubric (0 = not observed, 1 = rarely, 2 = occasionally, 3 = frequently, 4 = extensively) on 67 percent of the items and were within one category on the five-category

scale for 95 percent of the items. Similarly, a more recent study (Sterbinsky and Burke 2004) that included 16 paired comparisons in eight schools found that observer ratings were within one category for 91 percent of the targeted observations.

The Algebra I Quality Assessment is based on the Rubric for Student-Centered Activities. The rubric was developed by the Center for Research in Educational Policy (Lowther and Ross 2002) as an extension of the School Observation Measure to more closely evaluate the application quality (1 = limited application, 2 = somewhat limited application, 3 = somewhat strong application, 4=strong application) of seven student-centered learning activities (cooperative learning, project-based learning, higher-level questioning, experiential/hands-on learning, student independent inquiry/research, student discussion, and students as producers of knowledge using technology). Observers used operational definitions of the four levels of quality to designate ratings. The rubric reliability results were collected during16 paired comparisons of targeted observations in eight schools. Results indicate that there was an exact match of observer ratings for 63 percent of the comparisons, a match within one category for 27 percent of the comparisons, a match within two categories for 7 percent of the comparisons, and within three categories for 3 percent of the comparisons (Sterbinsky and Burke 2004).

Analytic Methods

The statistical analyses of the two key outcome measures took the same general approach, although the functional form of the estimating equations differed because the test score measure was a continuous variable and the math course variable was dichotomous. All impact estimates were generated using two-level hierarchical linear models (for test scores) or logit link models (for the 10th grade math enrollment) with standard errors that appropriately account for sample clustering. The robustness of the study findings were tested by sensitivity analyses. See appendix D for more details.

In all impact models, the treatment variable was a dichotomous variable indicating whether the student was enrolled in a treatment school or a control school. In the first primary confirmatory analysis for the models estimating impacts on test scores, the coefficient on the treatment variable is an estimate of the average treatment effect. The coefficient on the treatment variable in the model to estimate math course enrollment measures the marginal impact of being in the treatment group on the log of the probability of enrolling in an advanced math course. The estimate of the average treatment impact is calculated in four steps: (1) compute the estimated the probability of enrolling in advanced math for each student based on his or her individual characteristics and setting T to 1; (2) compute the probability for each student based on his or her characteristics and setting the treatment variable to 0; (3) computing the average probability assuming treatment equals 1 and assuming treatment equals 0; and (4) subtracting the estimated probability for the control group (treatment equals 0) from that for the treatment group (treatment equals 1).

Methodological issues

There are five methodological issues that warranted special attention. One is the treatment of missing data. A second relates to adjustments for multiple comparisons. A third pertains to weighting of student observations. A fourth relates to selection of covariates to be used in the analysis; and a fifth relates to the definition of the analysis sample.

Missing data

This study had relatively low levels of missing data for covariates. Thus, any of a number of methods for dealing with it would be acceptable (Puma, Olsen, Bell, and Price 2009). We chose to use casewise deletion of observations with missing outcome measures and dummy variable adjustments for missing values for covariates.

Casewise deletion was used for students missing outcome data. PLAN data were missing for students who were enrolled in a Kentucky public school in year 2 but did not take the PLAN (8 percent in the control group and 9 percent in the treatment group), students who were missing both PLAN data and year 2 enrollment data for unknown reasons (4 percent in the control group and 5 percent in the treatment group), and students lost to attrition who were no longer enrolled in a Kentucky public school in year 2 (2 percent each in the control and treatment groups). Outcome data for the grade 10 math course enrollment were missing for students who were no longer enrolled in a Kentucky public schools (1 percent each in the treatment and control group), and other students who were missing year 2 enrollment data for unknown reasons (6 percent in the control group and 7 percent in the treatment group).

Adjustments for multiple comparisons

No adjustments were made to reduce the risk of false discoveries due to multiple comparisons because math achievement and math course-taking are not the same outcome domain.

Weighting

The data were not weighted. The number of grade 9 algebra I students per school ranged from 20 to 373, and all grade 9 algebra I students in each school were included in the study. Larger schools received more weight in the hierarchical linear model analysis because they had more students in the level 1 sample.

Selection of covariates

As noted above, the basic models were estimated with all student- and school-level covariates except the student-level pretest scores. As part of the sensitivity analyses, separate models were estimated, including parameters associated with only student-level scale scores on the KCCT and EXPLORE pretests, student-level scale scores on the KCCT pretest, and student-level scale scores on the EXPLORE pretest. These models provided baseline estimates of impact of the intervention and helped determine whether the impact findings were sensitive to the model specification.

Then, the full models were estimated but with only one set of pretest variables at a time. Models were estimated excluding the variables for the EXPLORE school-level average pretest math scores and EXPLORE student-level deviation and excluding the

variables for the KCCT school-level average pretest math scores and KCCT student-level deviation. The reason is that the Kentucky Department of Education is interested in learning whether, controlling for the EXPLORE score in math, the intervention affects PLAN outcomes.

Definition of the study sample

As a sensitivity analysis, the full models were estimated on samples that were more restrictive than those used in the main analysis. One alternative sample excluded students in algebra 1 courses that began in eighth grade. Students in these part-year courses were exposed to the online courseware for only the fall semester of their ninth grade year. Moreover, their teachers had less experience using the materials and had participated in fewer school-year professional development sessions than would be the case for teachers of year-round courses. Additionally, for cohort I, students would have used only an older version of the student courseware that included errors at the beginning of the intervention year.

Impacts also were estimated on samples that excluded from cohort II the six duplicate schools that had been cohort 1 control schools and were rerandomized as part of cohort II.[26] The reason for this sensitivity analysis was to see if the findings were sensitive to excluding these six schools (strata 2).

Exploratory impact analyses

In addition to estimating impacts for the overall study sample, we also estimated impacts for subgroups defined by student and school-level characteristics. These analyses are considered exploratory and, thus, no corrections were made for multiple comparisons.

Student subgroups

Impacts for student subgroups were estimated by including variables that interacted treatment status with the student characteristics of interest. Separate analyses were conducted to estimate impacts for subgroups defined by gender and by enrollment cohort. The analytic models included school-level variables representing rural status, strata (for the gender subgroup analysis) or cohort (for the cohort analysis), and two school-level average pretest scores (KCCT and EXPLORE). In addition, they included student-level variables for free or reduced-price lunch status, gender, Individualized Education Plan status, underserved minority status, age, course level, and two measures of the student-level deviation from the school-level average pretest math score (KCCT and EXPLORE). The treatment variables included the binary variable indicating whether a student was enrolled in a treatment or control school, as well as this treatment variable interacted with the subgroup indicator of interest (male or cohort I). See appendix D for details on the model specifications.

The reason for estimating impacts for males and females is that others have found that males and females have different math problem-solving strategies and learning styles (Carr and Davis 2001; Friedman 1995; Geary, Saults, Liu, and Hoard 1999). These

[26] Nine students in the rerandomized schools repeated grade 9 and thus were included in the analysis samples for both cohort I and cohort II.

gender differences may have affected how students responded to the intervention. The reasons for examining cohort effects is that, during the first year of the intervention, the virtual school installed an early version of the student courseware that had multiple errors in the lessons and assignments, which were not fully corrected until the second semester. The impacts of the intervention may differ for students in cohort I because they had the incorrect version of the software for part of their exposure. Also, the instructional specialists did not have any experience with the hybrid algebra program in the first year of the program, which may have influenced the effectiveness of professional development.

School subgroups

In order to examine whether there were differences in the intervention's impacts between rural and non-rural schools, the models were modified to include a term that interacted the treatment variable and the rural status variable. Student characteristics were treated as fixed effects, and treatment effects were allowed to vary according to the locale of the school. It was expected that impacts might be different in rural and nonrural schools since rural areas are often economically depressed and geographically and socially isolated, thus, making it more difficult to attract highly qualified teaching candidates (McClure, Redfield, and Hammer 2003).

4. Implementing the Intervention

This chapter details how the intervention was implemented. It overviews hybrid algebra I in the classroom and explains the classroom setups, materials, and instructional practices. Next, it describes the components and materials of the professional development program and contrasts the intervention to the counterfactual conditions in the algebra I control group. It then addresses fidelity of implementation by discussing the development of fidelity rating scales and indicating the extent to which treatment schools were exposed to the intervention. Lastly, it provides the results from teacher surveys and classroom observations to describe variation in the classroom activities between the treatment and control schools.

Hybrid algebra I in the classroom

Hybrid algebra I was based on the implementation of research-based face-to-face and online learning strategies targeted to build student algebraic reasoning and skills through a standard three-part procedure for each lesson. Part 1 activated prior learning through warm-up activities such as "bell ringers," which consisted of short exercises completed by the students to review previously taught material. Part 2 introduced new knowledge and skills to be learned. Part 3 involved student reflection of learning through activities such as exit slips, in which students wrote responses to questions from the teacher that focused on synthesizing or summarizing new learning, making connections to the real-world context, or creating a bridge between new learning and learning to come in the next class period. The hybrid learning strategies aligned with the Kentucky algebra I curriculum standards and those of the National Council of Teachers of Mathematics (NCTM 1989, 1991, 2000), North American Council for Online Learning (NACOL 2007), and National Education Technology Standards for Teachers and Students (ISTE 1998, 2000).

The face-to-face instruction included whole-group, small-group, and individual activities. During whole-group time, teachers were instructed to use a computer connected to a projector to show the class new material. The teachers were encouraged to use concrete examples and questioning techniques to model elements of math literacy, such as reading, writing, speaking, and listening, to learn algebra concepts and skills. Small-group and individual activities included students constructing models of conceptual knowledge, representing their understanding in multiple ways, and monitoring their own thinking through reflection.

The online part of the course required students to use computers with high-speed Internet access at least two days per week (40 percent of class time) to complete algebra I courseware activities. The courseware provided students with teacher-selected lessons. Each lesson included visual representations, as well as audio and written descriptions, of algebraic concepts and problems to solve. The courseware enabled students to work at their own pace and provided immediate feedback on the correctness of each response. During this time, teachers served as facilitators of learning by providing individual assistance or small-group mini-lessons as needed.

Teacher preparation to implement hybrid algebra I consisted of a comprehensive one-year professional development program that used face-to-face and virtual/online formats, to give teachers a hands-on opportunity to learn the hybrid approach. The professional development focused on increasing teachers' conceptual understanding of algebraic reasoning and implementing research-based math pedagogy in a hybrid environment. Another goal was to create an online professional learning community of algebra I teachers.

Classroom setups

The Kentucky Virtual Schools' algebra I hybrid program suggests specific classroom configurations and requires student access to computers.

Classroom configuration

A standard grade 9 hybrid algebra I classroom organizes desks or tables so that students can sit in small groups for discussion and problem solving, or work individually with graphing calculators or manipulatives. This class arrangement also supports whole-group viewing of Internet-based algebra I instructional materials.

Access to computers

Teachers have computers with Internet access and a compatible projection device and screen that can be used during teacher-led instruction. Students have individual access to Internet-enabled computers at least two days per week, either through regular scheduling of the computer lab, laptops on carts, a one-to-one laptop program in which students have 24/7 access to computers, or assignment of the course to a classroom with one computer per student. Instructional specialists work with teachers to develop structures for moving students to computers efficiently in these different settings.

Materials

Implementing the intervention entails using a variety of technology tools and manipulatives to integrate hands-on activities with the instruction. One limitation is that there is no information on the extent to which the required and recommended technology tools and manipulatives were available to the treatment teachers and students.

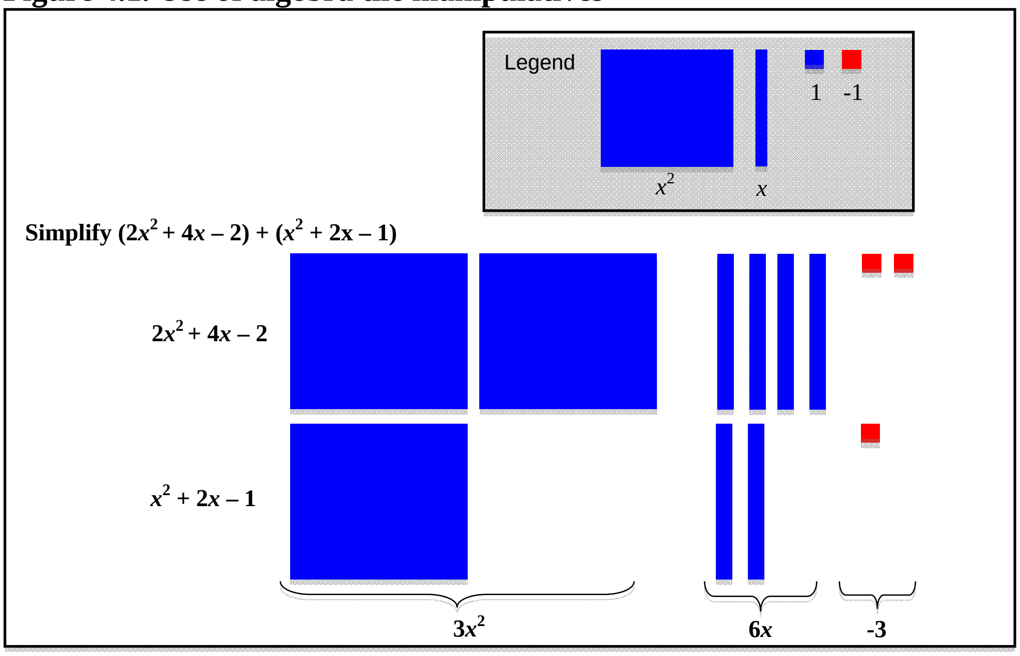
Technology tools

High-speed Internet access for teachers and students is required to integrate recommended web resources with hybrid algebra I instruction. An *Internet-connected projection device*, such as an interactive whiteboard, is required for use by teachers and students in the face-to-face setting to support such instructional activities as modeling and academic dialogue of algebraic concepts, use of virtual manipulatives, and vocabulary development through web-based activities. Also required are classroom sets of graphing calculators for student use and graphing calculator teacher tools for class demonstrations. The teacher tools require a device and/or software to project graphing calculator output to the class. Teacher and student use of spreadsheet and graphic software (MS Excel, Google Docs, etc.) is also recommended.

Manipulatives

Math manipulatives are "objects designed to represent explicitly and concretely mathematical ideas that are abstract. They have both visual and tactile appeal and can be manipulated by learners through hands-on experiences." (Moyer 2001, p. 176) In the hybrid algebra I intervention, four types of manipulatives are recommended: algebra tiles, geoboards, colored chips, and virtual manipulatives. The manipulatives enhance student understanding of algebraic concepts traditionally taught at the symbolic level, such as adding and subtracting polynomials (figure 4.1). Manipulatives are used during cooperative learning to engage students in algebraic discourse by giving students objects to think with and talk about.

Figure 4.1. Use of algebra tile manipulatives



Source: Researcher generated.

National Repository of Online Courses algebra I courseware

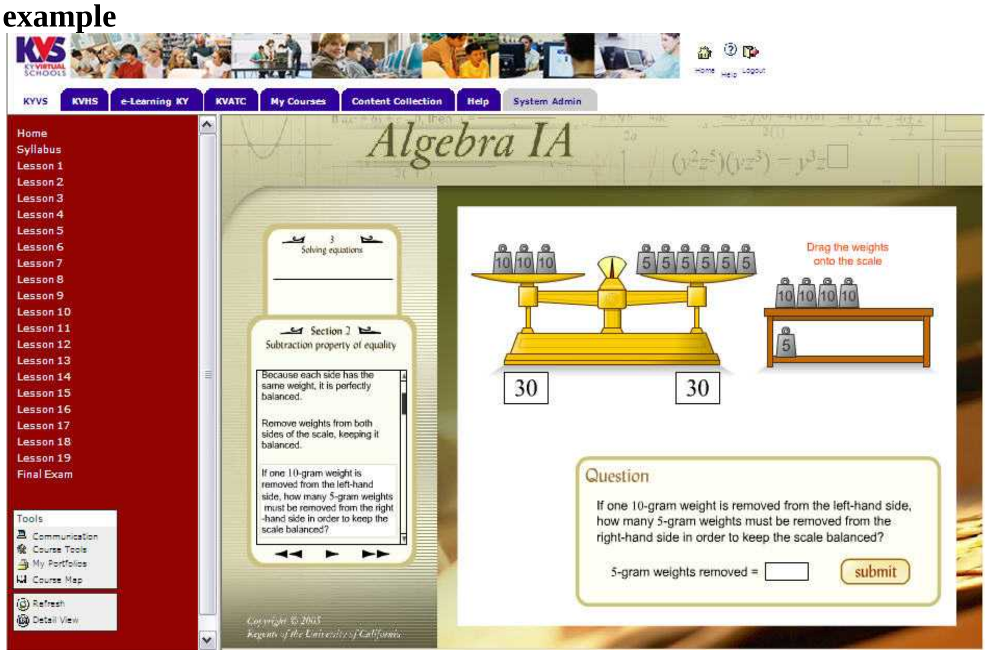
The Kentucky Virtual Schools/Kentucky Department of Education–selected courseware is an off-the-shelf product created by university faculty affiliated with the University of California College Prep Online and the Center for Digital Innovation at the University of California at Los Angeles. The courseware has gone through an external quality control protocol by the National Repository of Online Courses and Kentucky Department of Education curriculum specialists. One notable feature is its visual presentation of algebraic concepts alongside an audio description of the concept and written text that students can read with the audio. Kentucky Virtual Schools and the Kentucky Department of Education selected the courseware because it aligns closely with Kentucky algebra I standards; it is customizable, allowing educators to modify the content as needed; and program expansion is affordable due to the unlimited-use license associated with Kentucky Virtual Schools' membership with the National Repository of

Online Courses consortium.[27] These last two features have particular appeal for going to scale because the content can be tailored to meet the unique needs of different school systems and doing so would likely not be cost prohibitive.

For this study, Kentucky Virtual Schools registered all algebra I students in the treatment group for the National Repository of Online Courses algebra I courseware. Separate online class sections were set up to enroll students for each classroom teacher. The class sections were determined by the algebra I course schedule of each teacher.

Figure 4.2 is a sample of the repository's algebra I methodology, showing how to solve equations using the subtraction property of equality. The box on the left of the screen gives step-by-step instructions for the exercise. The main screen depicts a balance scale with three 10-gram weights on the left-hand side and six 5-gram weights on the right-hand side. A question box asks, "If one 10-gram weight is removed from the left-hand side, how many 5-gram weights must be removed from the right-hand side in order to keep the scale balanced?" Students can use the mouse to drag 5- or 10-gram weights on or off the scale to see how these changes affect either side of the scale. Once the student finds the number of 5-gram weights that need to be removed, the student enters the answer into a box at the bottom of the screen and receives feedback on the correctness of the response.

Figure 4.2. National Repository of Online Courses algebra I courseware lesson example



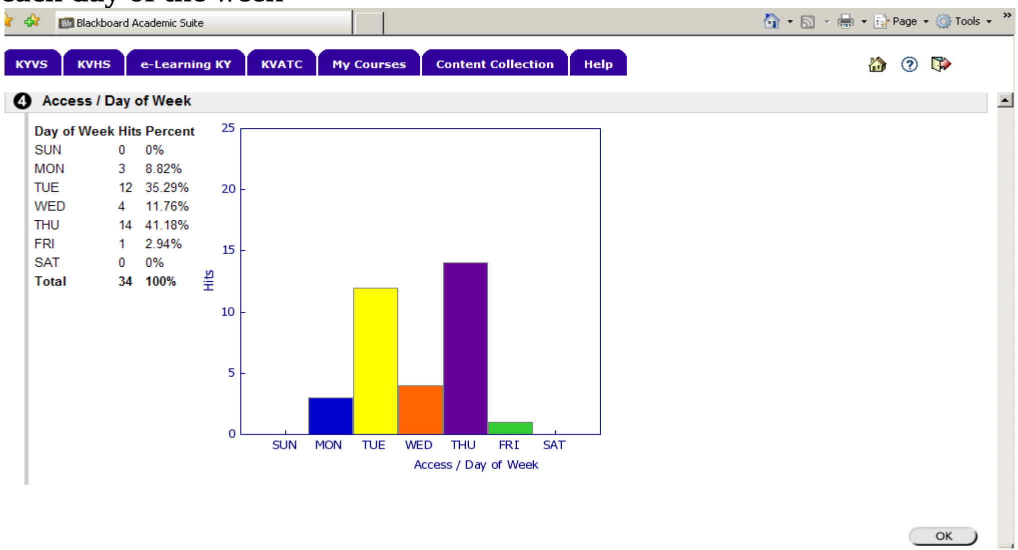
Source: National Repository of Online Courses.

[27] Kentucky Virtual Schools provides the Blackboard online learning system, access to the repository of online courses and instructional resources, registration services, a 24/7 help desk, initial teacher training, and hosting of the online environment. To obtain access to hybrid learning resources, schools/districts purchase accounts at $25 a student, which offer unlimited access to enrollment in online courses and community and instructional resources at no additional cost. Teacher and facilitator accounts are provided for free (see www.education.ky.gov/KYVSAssets/KYVS%20Blended%20Instruction.pdf).

Blackboard

Blackboard was the course-management system that delivered and supported the courseware for this study. Each hybrid algebra I course was assigned a designated online space on Blackboard, and only enrolled students had access (password-protected) to the courseware. Teachers and students simply logged onto the Kentucky Virtual Schools Blackboard site, with no need for specialized software. Hybrid algebra I teachers could monitor student progress through the online National Repository of Online Courses courseware with student usage statistics, such as number of logins, collected by the Blackboard system (figure 4.3). The Blackboard system captures and archives all course content without teachers having to reupload or recreate the materials for each school year. Thus, the same algebra I course content was used for cohorts I and II.

Figure 4.3. Student usage statistics on Blackboard with the number of hits during each day of the week



Source: Blackboard.

Earphones

Because the courseware includes audio instruction, students are required to wear earphones when listening to online instruction, to avoid interfering with other learners.

Technical requirements for the courseware and Blackboard

All schools participating in random assignment were required to complete a school administrator commitment form, requiring schools to verify the availability of technology that met the Kentucky Virtual Schools' Blackboard system requirements as listed at www.kyvs.org and as shown in table 4.1. Kentucky Virtual Schools also required high-speed Internet access for student use of online courseware.

Table 4.1. Technology requirements for hybrid algebra I courseware and Blackboard

System requirements	Microsoft Windows	Macintosh OS
Operating system:	Windows 2000 or PEOPLE	Mac OS 10.2, 10.3, or 10.4
Internet browser:	Microsoft Internet Explorer 6.0 or 7.0, Netscape 7.1 or 8.0, or Firefox 1.0	Netscape 7.1, Firefox 1.0, Apple Safari 1.0, 1.1, 1.2, or 2.0
Sound card and speakers	Required	Required
Web browser with Java enabled	Required	Required
Real One Player 2	Required	Required
Macromedia Flash player	Required	Required
Macromedia Shockwave player	Required	Required
High-speed Internet access	Required	Required

Source: Kentucky Virtual Schools' help desk website (www.kyvs.org).

Instructional practices

Implementing a hybrid algebra I class session as intended requires the teacher to plan four lesson components. The teacher sets learning goals for each lesson, conveys those goals to students, implements a strategy for reaching those goals, and assesses learning at the end of the class period. These steps are required for every lesson—whether or not students are using the online courseware on that day.

Learning goals for the lesson are to be displayed every day in a consistent location, such as an upper corner of a whiteboard, so that students know where to find the learning goals and other important information, such as homework assignments or the date of an upcoming exam. The teacher is to tell students the learning goals for the day early in the class period.

The hybrid algebra I program prescribes a standard three-part procedure for each lesson. Part 1 activates prior learning, part 2 introduces new knowledge and skills to be learned, and part 3 involves student reflection of learning.

Part 1: Each lesson brings with it an activity designed to activate prior knowledge associated with the lesson's learning goals. This might involve whole-class activities such as question-and-answer sessions, warm-up problems, discussions or activities, and individual or group activities, such as writing-to-learn tasks. The teacher may use writing-to-learn activities such as journal writing, or bell-ringer activities posted before class that students start when they enter the room and work on until the teacher rings a bell.

Part 2: Next, the teacher introduces new algebra I concepts and skills to be learned through face-to-face and/or online learning activities. For example, journal writing and other writing activities are used to reinforce newly introduced concepts. Also, on days when students work independently with online resources, the teacher introduces new concepts and skills that students will encounter in the online lesson (such as use of an online graph to show changes in the characteristics of a line as the coefficients change). In the online lessons, students can use a computer-based note-taking tool to track new learning, which promotes active rather than passive or minimal involvement during the lessons.

Transition to the online lesson involves moving to a computer lab scheduled for the class or using classroom computers. At the computer, students retrieve their earphones and log into Kentucky Virtual Schools to access the National Repository of Online Schools courseware, go to their lesson, and begin work using the note-taking device. Students may work alone or assist one another. The teacher monitors and facilitates student learning, helping individual students, a small group, or the whole class with a mini-lesson, as needed.

Part 3: During the final activity, students reflect on what they learned during the lesson. Reflection is intended to helps students process and retain the new information. Student reflection strategies include note-taking, exit slips, student journals, and small-group discussion. Teachers also use reflection as a formative assessment tool to identify what students learned and where they were confused, to inform the planning of their next lessons.

Hybrid algebra I professional development model

Treatment teachers participated in a comprehensive 12-month professional development program facilitated by two instructional specialists in math from the Collaborative for Teaching and Learning, a Louisville-based professional development provider. The instructional specialists were experienced professional development providers with backgrounds in math instruction at the secondary level and were trained by Kentucky Virtual Schools on using the primary online resources.

The hybrid algebra I professional development focuses on how to teach in a hybrid format and how to implement research-based practices. The professional development itself employs a hybrid approach that includes face-to-face teacher training using online materials and fully online formats.

The two primary reasons for using hybrid techniques for the teacher training are that it helps teachers understand the challenges faced by their students when learning via a hybrid format, and it increases teachers' frequency of learning-by-doing interactions while building a professional learning community.

The professional development focuses on balancing understanding of algebra I content with development of pedagogy specific to algebra I (National Standards Development Council 2001) as applied in a hybrid instructional environment. Focusing professional development on content, both subject knowledge and how students learn the subject, has been associated in several studies with effective professional development (Banilower, Heck, and Weiss 2005; Cohen 1990; Cohen and Hill 2001; Garet, Porter, Desimone, Birman, and Yoon 2001; Kennedy 1998; Smith et al. 2007).

In addition, the professional development incorporates design features—content focus, active learning, coherence, duration, collective participation—that have been shown in empirical literature to be associated with improved instructional practices and, with more limited evidence, an increase in student achievement. In particular, it uses active learning: the teachers observe the instructors and their techniques, have two opportunities for their classrooms to be observed on site by instructional specialists during the school year, and continually engage in feedback and collaboration through

online discussions. All these components are associated with increased professional development effectiveness (Banilower and Shimkus 2004; Borko 2004; Carey and Fechtling 1997; Darling-Hammond 1997; Lieberman 1996). Desimone (2009) provides an overview of this research base.

Further, the length and intensity of professional development, facilitated by a mix of face-to-face and online venues for learning, are consistent with the evidence on the importance of these features in professional development programs associated with changes in teacher practices (Desimone 2009). Finally, the initial face-to-face sessions, online discussions, and enrollment of algebra teachers in each participating school help build a learning community, which involves interaction and communication among teachers, both shown to be influential in teacher learning (Banilower and Shimkus 2004; Borko 2004; Desimone 2003; Fullan 1991; Guskey 1994; Little 1993; Loucks-Horsley, Hewson, Love, and Stiles 1998; Rosenholtz 1989).

Professional development components

The two instructional specialists work with the teachers, who are divided into two approximately equal-sized groups determined by geographic proximity to the face-to-face training sessions and that vary in composition, based on teaching schedules or other commitments for the online sessions. The yearlong professional development occurs in five segments: initial face-to-face training, summer online training, end-of-summer face-to-face session school-year online sessions, and two school-year site visits to each school. A brief overview of each segment is provided below, and a timeline with additional information is presented in appendix F.

1) Goals of first face-to-face sessions

The primary goals of the initial face-to-face sessions are to introduce teachers to the program. For teachers in the research study, this session also introduces them to their roles and responsibilities as participants, so the initial one-day session begins with an orientation on the research objectives and design. Attention then turns to the hybrid algebra I program itself. Each teacher receives a binder with materials (see below) that the instructional specialists examine with the group. In addition, the intervention teachers are introduced to the National Repository of Online Courses online courseware for algebra I student instruction, the professional development courseware (Spotlight on Algebra I), and Blackboard. A two-day face-to-face training session follows the orientation, leading teachers through four spotlight sessions, as well as Horizon Wimba, a communication system used by the Kentucky Department of Education for online conferencing. The three-day program immerses teachers in learning activities and begins to create a learning community extending throughout the program.

2) Goals of summer online training

The primary goal of the summer online training is to strengthen and improve algebra I teaching skills and practices. During the five weeks of training, teachers complete Spotlight on Algebra I activities and engage in weekly online discussions with other participating teachers, all guided by the instructional specialists.

3) Goals of end-of-summer face-to-face session

The end-of-summer face-to-face session prepares teachers to register students in the online course and review guidelines for teaching hybrid algebra I with the National Repository of Online Courses courseware. The session also reviews the use of Blackboard for course management, the professional development schedule, and procedures for participating in the online activities during the school year.

4) Goals of school-year online sessions

The key goals of the school-year online professional development sessions are for teachers to share teaching strategies and resources, discuss planning instruction, analyze student work, and share formative and summative evaluation strategies to move learning forward. The professional development continues throughout the school year to provide ongoing embedded support through monthly online sessions with the community of learners.

5) Goals of school-site visits

An instructional specialist visits each treatment teacher in the fall semester to observe instruction, answer questions, and provide individual feedback on instructional practice. A second visit follows early in the spring semester. The site visits are intended to help teachers implement the hybrid approach as designed. The visits also inform the instructional specialists on common mistakes and challenges facing teachers as they implement the hybrid approach. Instructional specialists use this formative feedback when planning discussion topics for the monthly online sessions.

Professional development materials

There are three primary resources for professional development: the hybrid algebra I professional development teacher binder, the Spotlight on Algebra I online curriculum, and the Horizon Wimba communication software.

Teacher binder

Teachers receive a binder designed by the Collaborative for Teaching and Learning's math instructional specialists. The binder includes support materials for each professional development session and space for teachers to add completed work and notes. It also includes a copy of Kentucky's standards for high school math and a chart showing how the student courseware aligns with those standards. Gaps in content of the courseware are highlighted to call teachers' attention to required content not supported by the courseware. National and state technology standards and Internet addresses for online math instructional resources are also included.

Spotlight on Algebra I

The primary instructional component of professional development is Spotlight on Algebra I (figure 4.4). The courseware, developed by the Southern Region Education Board with additional funding from the AT&T Foundation, has been available to teachers in the Southern Region Education Board region since 2001. The Spotlight on Algebra I lessons use three strategies to teach algebra concepts:

- *Thinking about content.* Provides an overview of the key algebra concepts covered.
- *How do I teach this?* Presents instructional strategies that engage students in using multiple ways to solve algebra problems. Traditional approaches include use of pictures, diagrams, charts, tables, and graphs in addition to symbolic notation. Nontraditional strategies engage students in critical-thinking activities involving reading, writing, talking, and listening to help students translate mathematical ideas to real-life situations and build a deeper understanding of math.
- *Helping students.* Has teachers analyze student work not only for correct answers but also for the kinds of obstacles students encounter when trying to solve problems. These activities emphasize the ultimate goal of improving student learning.

Figure 4.4. Spotlight on Algebra I home page

Source: Spotlight on Algebra I.

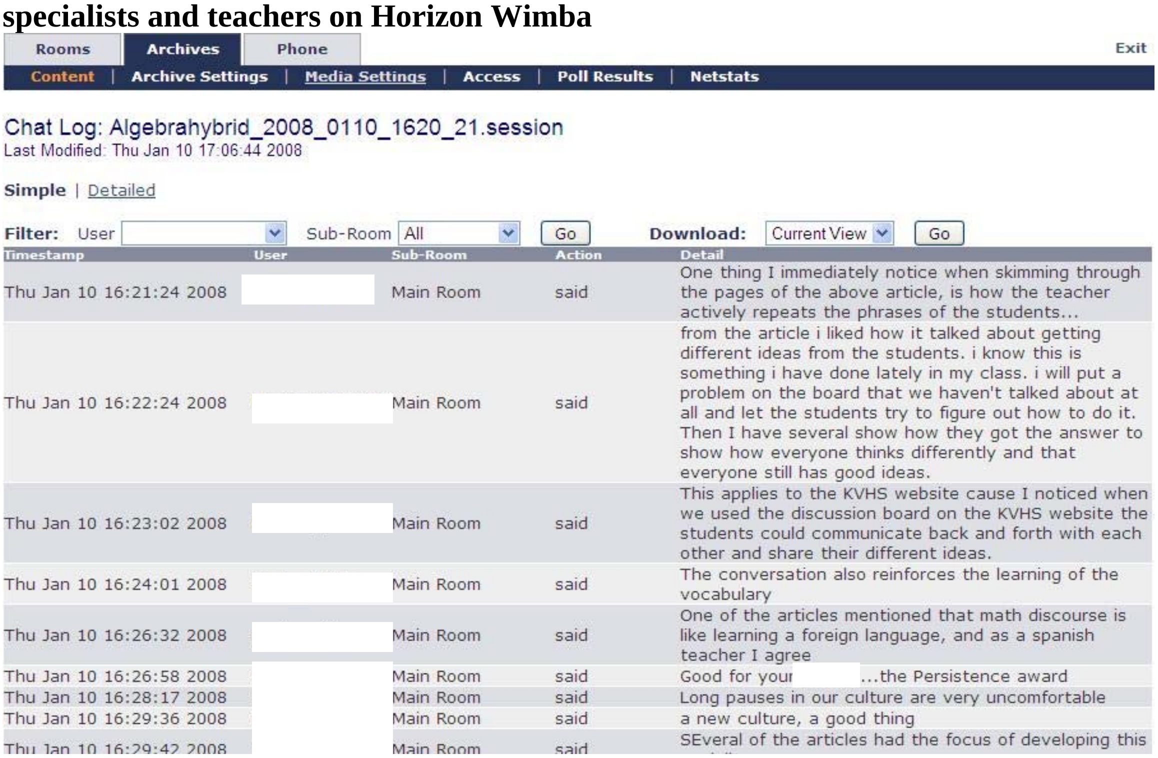
National Repository of Online Courses courseware

The National Repository of Online Courses courseware is also used during professional development to help teachers plan and deliver instruction.

Horizon Wimba

Horizon Wimba is an online system that permits voice communications among participants, instant messaging, and visual display of instructional materials. It is used to support communication among the instructional specialists and teachers. During the school year, instructional specialists use Horizon Wimba to lead monthly online discussions to help teachers build context for instruction, select activities well suited to key mathematical concepts, and discuss classroom plans and experiences (figure 4.5). Multiple sections of these online sessions are held to accommodate a wide range of teaching schedules. Teachers receive a schedule of upcoming sessions and are required to participate in at least one session per month.

Figure 4.5. Sample chat log from an online discussion session between instructional specialists and teachers on Horizon Wimba



Rooms | Archives | Phone | Exit

Content | Archive Settings | Media Settings | Access | Poll Results | Netstats

Chat Log: Algebrahybrid_2008_0110_1620_21.session
Last Modified: Thu Jan 10 17:06:44 2008

Simple | Detailed

Filter: User | Sub-Room All | Go | Download: Current View | Go

Timestamp	User	Sub-Room	Action	Detail
Thu Jan 10 16:21:24 2008		Main Room	said	One thing I immediately notice when skimming through the pages of the above article, is how the teacher actively repeats the phrases of the students...
Thu Jan 10 16:22:24 2008		Main Room	said	from the article i liked how it talked about getting different ideas from the students. i know this is something i have done lately in my class. i will put a problem on the board that we haven't talked about at all and let the students try to figure out how to do it. Then I have several show how they got the answer to show how everyone thinks differently and that everyone still has good ideas.
Thu Jan 10 16:23:02 2008		Main Room	said	This applies to the KVHS website cause I noticed when we used the discussion board on the KVHS website the students could communicate back and forth with each other and share their different ideas.
Thu Jan 10 16:24:01 2008		Main Room	said	The conversation also reinforces the learning of the vocabulary
Thu Jan 10 16:26:32 2008		Main Room	said	One of the articles mentioned that math discourse is like learning a foreign language, and as a spanish teacher I agree
Thu Jan 10 16:26:58 2008		Main Room	said	Good for your ...the Persistence award
Thu Jan 10 16:28:17 2008		Main Room	said	Long pauses in our culture are very uncomfortable
Thu Jan 10 16:29:36 2008		Main Room	said	a new culture, a good thing
Thu Jan 10 16:29:42 2008		Main Room	said	SEveral of the articles had the focus of developing this

Source: Horizon Wimba.

Counterfactual conditions: algebra I control group

Study application materials indicated that the algebra I teachers in control schools would teach algebra I in a "business as usual" manner. Follow-up discussions with eligible schools clarified that control schools would not be asked to change their teaching plans or practices to participate in the study.

When this study began, Internet access was universal in Kentucky schools; use of technology for student learning was expanding rapidly and learning resources were becoming more available through Internet searches. As noted, Spotlight on Algebra I materials were provided for free to Kentucky educators. However, because Kentucky Virtual Schools had not offered a facilitated version of it in several years, and because of the length of this course, few Kentucky teachers were expected to have completed the professional development program on their own. To minimize "leakage" of the professional development experience to teachers in control schools, Kentucky Virtual Schools agreed to limit access to Spotlight on Algebra I to treatment teachers for the duration of the study.

Fidelity analysis

Multiple components are involved in the implementing hybrid algebra I intervention. First, it requires that teachers complete all components of the professional development training. Second, it requires use of research-based practices 60 percent of instructional time and student use of algebra I courseware 40 percent of the time, via one-to-one access to computers with reliable high-speed Internet connections. The fidelity analysis sheds some light on the degree to which these treatment conditions were implemented in a real-world setting.

Researchers were not involved in the implementation, though they monitored its progress through regularly scheduled conversations with the instructional specialists. Thus, some hindrances to implementation were expected, including technical problems such as Internet connectivity issues and problems with Kentucky Virtual Schools' courseware delivery (for example, server problems and teachers unable to reach the help desk).

An unanticipated concern regarding use of the online courseware is noteworthy. The student courseware was the latest version available at the time the intervention began. School systems using the courseware may stream it directly from the provider or obtain DVDs that can install the courseware on the learning platform used by the school system. In this study, the content from DVDs for both the student courseware and teacher professional development materials were placed onto a Blackboard learning platform, an approach that provides teachers access to a variety of online course management tools, and enables educators to time the installation of updates to courseware. In this case, however, the virtual school installed an early version of the student courseware by mistake. The installed version had not been through careful review, and student exercises had numerous errors. Complaints from teachers and instructional specialists led to an inquiry by the research team, the discovery of the problem, and the problem's resolution. Updated courseware was installed at the start of the second semester of the first cohort.

Additional issues, including turnover of treatment teachers during the school year for reasons unrelated to the intervention, such as long-term illness or teachers not returning to their positions, led to the late hiring of teachers. These hiring delays required that new teachers be trained with a condensed professional development model if they agreed to participate in the intervention. Another challenge in implementing the intervention was achieving school-level participation. In the school year before the intervention,

superintendents and principals signed agreements to include all algebra I teachers, and teachers in those schools signed an application form agreeing to participate. Teachers hired after the school year before the intervention, when initial agreements were signed, were asked to sign an agreement when the instructional specialists or researchers were notified of the new hire. Despite these agreements, 20 of the 63 teachers (31.7 percent) in treatment schools did not participate in any components of the intervention. Of those, 6 did not participate because their school withdrew from the study after randomization, and 14 attended no professional development sessions and did not use the online student courseware.[28] The remaining 43 treatment teachers (68.3 percent) were at least partly exposed to the intervention. Scores for fidelity of implementation were calculated for all treatment teachers to identify the number and percentage of students by implementation level.

Assessment of fidelity of implementation was based on the two foci of the hybrid algebra I course: professional development and hybrid instruction. All fidelity of implementation data collection instruments and scales were developed before actual data collection. The researchers created a common rating scale of high, moderate, low, and none to rate the fidelity of implementation. They also developed operational definitions for the scales based on the School Observation Measure and Algebra I Quality Assessment rubrics described in chapter 3.

Data on professional development were collected by the instructional specialists with the hybrid algebra I implementation fidelity record sheet. Fidelity ratings were collaboratively determined by the two instructional specialists to ensure inter-observer reliability. Data on the fidelity of hybrid instruction were collected from Blackboard records of student use of the algebra I courseware, and observational data were collected by instructional specialists with the fidelity of implementation instrument and by external site researchers using two instruments: the School Observation Measure and the Algebra I Quality Assessment. Descriptive statistics for the implementation fidelity ratings and an overall fidelity indicator were reported. All fidelity ratings and scales were specified before examining the data.

Fidelity of professional development

Fidelity of the professional development implementation was assessed with data on treatment teacher attendance, effort, and engagement in the summer professional development and the monthly online professional development sessions conducted during the school year for all treatment teachers.

Assessment of attendance

Records of teacher professional development participation were used to rate each treatment teacher's attendance for the summer (nine sessions) and the school year (nine sessions), using the four-point professional development attendance rating scale (table

[28] The list of algebra I teachers provided by the treatment schools was incomplete as all courses that provide algebra I credit were not included (nontraditional algebra I courses designed for students repeating grade 9 or for special education students). Once course enrollment records were received from the Kentucky Department of Education late in the semester, 14 teachers who should have participated in the intervention were identified.

4.2). The findings show that 59.1 percent of students in the treatment group had teachers who attended six to nine summer professional development sessions and 51.2 percent had teachers who attended six to nine school year professional development sessions, indicating high to moderate professional development attendance. Of the treatment group students, 15.9 percent had teachers who did not participate in summer professional development and 16.6 percent had teachers who did not participate in the academic year professional development online sessions.

Although the impact of use of the erroneous student courseware is unknown, if teachers left the program because of frustration with the courseware, the effect would be evident in increased attendance rates in the professional development program during the school year in cohort II, compared with cohort I. Table 4.2C compares teacher attendance rates by cohort, showing that this was not the case, as 31.3 percent of teachers in cohort I had high attendance, compared with 32.3 percent in cohort II. Participation in professional development is one aspect of program fidelity, and it is not clear whether or how the software glitches would influence this measure.

Table 4.2. Fidelity rating scale for professional development attendance

A. Rating scale for summer professional development attendance

n = 63 treatment teachers*

Rating scale	Rating	Description	Treatment teachers	Students of treatment teachers
3	High	8 to 9 sessions (89 percent to 100 percent)	22 (34.9 percent)	1,484 (43.7 percent)
2	Moderate	6 to 7 sessions (67 percent to 78 percent)	8 (12.7 percent)	522 (15.4 percent)
1	Low	1 to 5 session(s) (11 percent to 55 percent)	13 (20.6 percent)	848 (25.0 percent)
0	None	Did not attend any sessions	20 (31.2 percent)	541 (15.9 percent)

* Includes 20 teachers who did not participate in any components of the intervention.
Source: Authors' analysis of summer professional development attendance records (2007/08 and 2008/09).

B. Rating scale for school-year professional development attendance

n = 63 treatment teachers*

Rating scale	Rating	Description	Treatment teachers	Students of treatment teachers
3	High	8 to 9 sessions (89 percent to 100 percent)	20 (31.8 percent)	1,525 (44.9 percent)
2	Moderate	6 to 7 sessions (67 percent to 78 percent)	7 (11.1 percent)	213 (6.3 percent)
1	Low	1 to 5 session(s) (11 percent to 55 percent)	15 (23.8 percent)	1095 (32.3 percent)
0	None	Did not attend any sessions	21 (33.3 percent)	562 (16.6 percent)

* Includes 20 teachers who did not participate in any components of the intervention.
Source: Authors' analysis of academic-year professional development attendance records (2007/08 and 2008/09).

C. Rating scale for school-year professional development attendance, by cohort

n = 63 treatment teachers*

Rating scale	Rating	Description	Treatment teachers: Cohort I *n* = 32	Treatment teachers: Cohort II *n* = 31
3	High	8 to 9 sessions (89 percent to 100 percent)	10 (31.3 percent)	10 (32.3 percent)
2	Moderate	6 to 7 sessions (67 percent to 78 percent)	3 (9.4 percent)	4 (12.9 percent)
1	Low	1 to 5 session(s) (11 percent to 55 percent)	7 (21.9 percent)	8 (25.8 percent)
0	None	Did not attend any sessions	12 (37.5 percent)	9 (29.0 percent)

* Includes 20 teachers who did not participate in any components of the intervention.
Source: Authors' analysis of academic-year professional development attendance records (2007/08 and 2008/09).

Assessment of effort and engagement

The instructional specialists rated individual treatment teachers' effort and engagement during the summer and again for the school-year professional development with the effort and engagement rating scale (table 4.3). Teachers who did not participate in any professional development sessions received a rating of "none." Effort and engagement results show that 59.8 percent of the treatment group students had teachers who were rated as having high to moderate effort and engagement during the summer professional development sessions, whereas 47.7 percent had teachers who received high to moderate ratings for effort and engagement during school-year professional development.

Table 4.3. Rating scale for teacher effort and engagement during professional development

n = 63 treatment teachers*

Rating scale	Rating	Description	Treatment teachers: Summer	Treatment teachers: School year	Students of treatment teachers: Summer	Students of treatment teachers: School year
3	High	Fully completed all assignments, asked multiple questions, frequently contributed meaningful responses to the discussions, and regularly shared high-quality resources.	15 (23.8 percent)	14 (22.2 percent)	1119 (33.0 percent)	1185 (34.9 percent)
2	Moderate	Partially completed assignments, asked some questions, occasionally contributed to discussions or contributed moderate responses, shared some resources.	15 (23.8 percent)	9 (14.2 percent)	909 (26.8 percent)	436 (12.8 percent)
1	Low	Completed few assignments, rarely asked questions, infrequently contributed to discussions or contributed minimal responses, shared few if any resources.	12 (19.1 percent)	19 (30.2 percent)	804 (23.7 percent)	1212 (35.7 percent)
0	None	Did not participate in any professional development sessions.	21 (33.3 percent)	21 (33.3 percent)	563 (16.6 percent)	562 (16.6 percent)

* Includes 20 teachers who did not participate in any components of the intervention.
Source: Authors' analysis of professional development effort and engagement rating forms completed by instructional specialists (2007/08 and 2008/09).

Fidelity of hybrid instruction

Two data sources were used to assess the fidelity of implementation of the hybrid approach to algebra I instruction: student use of the courseware and classroom practices observational data collected by instructional specialists and external site researchers. For the first data source, fidelity ratings were assigned to all treatment teachers. Fidelity ratings based on observational data were only available for teachers who attempted to implement the intervention and were observed. The conclusions that can be drawn about the teachers' instructional strategies from the observations are limited because not all teachers were observed (the response rates were 68.3 percent for observations by instructional specialists and 60.3 percent for observations by external site researchers), most teachers were only observed on one day, and the reliability indices were low for some subscales.

Use of algebra I courseware

The data available for fidelity of implementation included the number of student connections per week from the electronic archives of the courseware, but not the actual amount of class time spent using the courseware. As a result, whether teachers met the stated expectations for instruction time to be split at 60 percent for face-to-face instruction and 40 percent for individual student use on the courseware cannot be calculated. The implementation fidelity for the use of the courseware was estimated by examining whether students made connections with it on an average of at least two of five days of the week (2/5 = 40 percent) from February 28 to April 30. The extent to which students made a connection with the courseware was not the same construction as a 60-40 split in instructional time, but it is the only figure available on the use of the courseware (table 4.4). Researchers collected student courseware usage data in June at the end the intervention year. Resulting data indicated that 35.3 percent of students had teachers with high (n = 431, 12.7 percent) to moderate (n = 768, 22.6 percent) use and 38.7 percent of students (n = 1,314) had teachers with low ratings of less than one student connection per month.[29]

[29] As a further check of the impact of the incorrect student courseware on use of the intervention, these results were disaggregated by cohort but no differences in rates of moderate to high use of the student courseware were found.

Table 4.4. Rating scale for student use of the Kentucky Virtual Schools' online algebra I materials
n = 63 treatment teachers*

Rating scale	Rating	Description	Treatment teachers	Students of treatment teachers
3	High	The average number of student connections per week is two or more.	7 (11.1 percent)	431 (12.7 percent)
2	Moderate	The average number of student connections per week is from one to less than two.	13 (20.6 percent)	768 (22.6 percent)
1	Low	The average number of student connections per week is less than one, but the average connection per student during one month is at least one.	10 (15.9 percent)	882 (26.0 percent)
0	None	The average number of student connections per month is from zero to less than one.	33 (52.4 percent)	1,314 (38.7 percent)

* Includes 20 teachers who did not participate in any components of the intervention.
Source: Authors' analysis of Kentucky Virtual Schools' algebra I courseware student usage data (2007/08 and 2008/09).

Instructional practices

The rating scale for use of recommended algebra I practices was derived from average ratings recorded by the instructional specialists during site visits with 43 treatment teachers and the average School Observation Measure and Algebra I Quality Assessment ratings recorded by external site researchers during class observations of 38 of those 43 teachers who at least attempted to implement the intervention. The instructional practices rating scale descriptions are based on the School Observation Measure and Algebra I Quality Assessment rubrics and were developed before the intervention was implemented.

As noted, the instructional specialists conducted regularly scheduled on-site meetings with treatment teachers to help them implement the intervention and to observe teacher use of the hybrid approach. Instructional specialists used the rating scale in table 4.5A to record the degree (high, moderate, low, none) to which each of the 43 treatment teachers with site visits was observed implementing recommended hybrid algebra I instructional practices. Data from instructional specialist observations indicate that 70.2 percent of students with site visits had teachers with high (*n* = 817, 28.6 percent) to moderate (*n* = 1,188, 41.6 percent) ratings for use of the recommended practices. However, these results do not include 20 teachers who did not attempt to implement the intervention.

Observations by the external site researchers revealed that 92.7 percent of students with observation data had teachers with high to moderate ratings (*n* = 2,440, 92.7 percent) for use of the recommended practices. These ratings exclude 20 teachers who did not attempt to implement the intervention and five additional teachers who were not observed by the external site researchers.

Table 4.5. Rating scale for use of recommended hybrid algebra I instructional practices

A. Instructional specialist

n = 43 treatment teachers who attempted to implement the intervention

Rating scale	Rating	Description	Treatment teachers	Students of treatment teachers
3	High	Recommended hybrid instructional practices receive substantive time (frequent to extensive use) and are a prevalent component of the teaching/learning during class.	10 (23.3 percent)	817 (28.6 percent)
2	Moderate	Recommended hybrid instructional practices receive moderate time (occasional use) and are a somewhat prevalent component of the teaching/learning during class.	18 (41.9 percent)	1,188 (41.6 percent)
1 to 0*	Low to None	Low: Recommended hybrid instructional practices receive isolated or little time (rare use) and are not a prevalent component of the teaching/learning during class. None: No use of recommended hybrid instructional practices.	15 (35.0 percent)	849 (29.8 percent)

*The number of teachers in each category of “low” to “none” is too few to report due to potential disclosure risk.
Source: Authors’ analysis of observation rating forms completed by instructional specialists (2007/08 and 2008/09).

B. External site researchers

n = 38 treatment teachers who attempted to implement the intervention and who were observed by external site researchers

Rating scale	Rating	Description	Treatment teachers	Students of treatment teachers
3 to 2*	High to moderate	High: Recommended hybrid instructional practices receive substantive time (frequent to extensive use) and are primarily rated as high quality. Moderate: Recommended hybrid instructional practices receive moderate time (occasional use) and are primarily rated as moderate quality.	34 (89.5 percent)	2,440 (92.7 percent)
1	Low	Recommended hybrid instructional practices receive isolated or little time (rare use) and are primarily rated as low quality.	4 (12.0 percent)	191 (7.3 percent)

*The number of teachers in each category of “high” to “moderate” is too few to report due to potential disclosure risk.
Source: Authors’ analysis of the School Observation Measure and the Algebra I Quality Assessment (2007/08 and 2008/09).

Overall indicator for fidelity of implementation

Information from tables 4.2–4.5 were used to create a global fidelity score that would be an overall indicator of the fidelity with which the hybrid algebra I intervention was implemented. The global fidelity scores, shown in table 4.6, were calculated as the sum of the seven scores for each treatment teacher. For teachers without a site visit or observation data, the global fidelity scores were developed by calculating the average score from tables 4.1–4.3 and multiplying by seven to create a global scale. The global fidelity of implementation rating scale of high, moderate, low, and "no meaningful use" was developed before data collection. It is important to note that the global fidelity score gives equal weight to each of the seven measured aspects of implementation. Conceptually, the global fidelity score assumes that each of the measurements are equally important, which may not be the case. For example, student use of courseware has a single metric and receives one-seventh weight. Teacher attendance and effort and engagement in professional development in the summer and school-year programs are captured by four separate measures and get four-sevenths weight.

Moderate Exposure to Treatment. Students in treatment classes are defined as having received a moderate amount and quality of treatment if the following criteria have been met:

- The teacher participated in at least two-thirds (6 of 9) of the professional development summer sessions.
- The teacher participated in at least two-thirds (6 of 9) of the professional development sessions in the school year.
- The teacher manifested at least moderate effort and engagement in the professional development during the summer.
- The teacher manifested at least moderate effort and engagement in the professional development during the school year.
- Student connections to the courseware averaged one or more per week over the period of February 28 to April 30.
- The instructional specialist observed at least moderate use of recommended hybrid instructional practices in the classroom.
- The site researchers observed at least moderate use of recommended hybrid instructional practices in the classroom.

For example, a teacher receiving a score of moderate (2) for each item would receive a global fidelity score of 14, which corresponds to a global score rating of moderate. However, since the global fidelity scores were based on a simple sum of the items, not all the criteria must be met before the treatment is assessed as adequate. For example, a teacher receiving a high (3) rating for summer professional development, a low (1) rating for school-year professional development, and moderate (2) ratings for all other items would also have a global fidelity score of 14.

Results. The fidelity of implementation analysis included data from the 63 treatment teachers on their attendance at summer and school year professional development sessions, effort and engagement during the professional development sessions, student use of online materials, and teacher use of recommended algebra I instructional practices as observed by instructional specialists and external researchers. Overall, evidence of implementation fidelity indicated that 56.5 percent of the treatment group students had teachers with high (*n* = 700, 20.6 percent) to moderate (*n* = 1,220, 35.9 percent) global fidelity score ratings, whereas, 27.5 percent (*n* = 934) of students had teachers who received low global fidelity ratings and 15.9 percent (*n* = 541) of the treatment group students had teachers who received a global score of "no meaningful use" (see table 4.6).

Table 4.6. Global fidelity scores
n = 63 treatment teachers*

Global score rating	Scale: Greater than or equal to	Scale: Less than or equal to	Treatment teachers	Students of treatment teachers
High	17	21	9 (14.3 percent)	700 (20.6 percent)
Moderate	12	16	18 (28.5 percent)	1,220 (35.9 percent)
Low	7	11	16 (25.4 percent)	934 (27.5 percent)
No meaningful use	0	6	20 (31.8 percent)	541 (15.9 percent)

* Includes 20 teachers who did not participate in any components of the intervention.
Source: Authors' analysis of 2007/08 and 2008/09 data for the summer and school-year professional development attendance records, professional development effort and engagement rating forms completed by the instructional specialists, Kentucky Virtual Schools' algebra I student usage data, observation rating forms completed by the instructional specialists for the School Observation Measure, and the Algebra I Quality Assessment.

Treatment-control contrasts

Data to contrast the treatment and control conditions were collected with teacher surveys and external site researcher classroom observations. The findings from these instruments (see chapter 3) are presented below. These results are descriptive. They contrast the treatment and control groups and should not be used to infer causal impacts. Tests of statistical significance were conducted to identify areas of difference between the two comparison groups; however, some of the significant differences may be false discoveries occurring by chance, due to the number of items examined. No adjustments were made for multiple comparisons since the findings are not intended to test the impact of the intervention.

Teacher survey results

The Hybrid Teacher Questionnaire and the Control Teacher Questionnaire were used to determine teacher perceptions on the instructional approach they used for algebra I during the study: the hybrid approach for treatment teachers or district algebra I curriculum for control teachers. The return rate for the Hybrid Teacher Questionnaire was 68 percent (43 of 63 treatment teachers), and the return rate for the Control Teacher Questionnaire was 70 percent (47 of the 67 control teachers).[30] The survey was not administered to 12 algebra 1 teachers in treatment schools who were not identified until

[30] Two of the 43 treatment teachers taught separate fall and spring algebra I courses and completed separate surveys for each course, so the number of treatment teacher surveys is 45.

course enrollment records were received from the Kentucky Department of Education late in the school year and 9 teachers who taught at schools that were randomized into the control group but did not participate in the data collections. For all other teachers, follow-up emails and phone calls were sent to non-responders in an effort to increase response rate.

The Wilcoxon-Mann-Whitney test was used to determine if there were significant differences in teacher attitudes toward using the hybrid algebra I approach (treatment) or the district's algebra I curriculum (control; table 4.7). First was an analysis of treatment and control teacher responses to the 15 items representing the three dimensions noted in the survey descriptions: Impact of hybrid/district algebra I on instruction (items 2, 3, 9, 11, 15); readiness to implement hybrid/district algebra I (items 7, 8, 10, 12, 13); and impact of hybrid/district algebra I on students (items 1, 4, 5, 6, 16). Item 14, "I can readily obtain answers to questions regarding implementation of the hybrid algebra I approach" was not associated with the three scales. Second, two composite scores were created to represent self-reported frequency of teacher use of algebra I strategies (ask *why* and *what if* questions; use number lines, graphs, or diagrams to explain algebra; and use a computer to teach algebra) and student use of algebra I strategies (work in groups; write to explain algebra; talk to explain algebra; use things like algebra tiles or blocks; use activities such as "guess and check," estimating, or drawing; use graphing calculators; use computers to learn algebra; and use exit slips). Third, a stand-alone question asked teachers to indicate whether they would like to use the same approach (hybrid or district curriculum) to teach algebra I again. All the dimension and composite scores were defined before examining the data and were calculated as an unweighted average of the corresponding survey items. See appendix G for detailed findings of perceptions for treatment and control teachers for each item, as recorded on the two questionnaires.

The teacher survey results revealed similar treatment and control teacher perceptions of the algebra I approach they used for two of the three dimensions. Specifically, there were no statistically significant differences in treatment and control teacher perceptions of their readiness to teach their assigned algebra I approach or the impact of the approach on instruction. The third dimension indicated that the treatment teachers felt their approach had less of an impact on students than control teachers (difference = –0.23, effect size = –0.34, $p = .04$).

Next, the self-reported frequency of use of algebra I strategies by teachers and students were examined. There was no statistically significant difference between the treatment and control groups in the use of algebra I strategies by teachers. However, the results revealed a statistically significant difference favoring treatment teachers' (mean = 3.28) over control teachers' (mean = 2.90) self-reported frequency of "student use of algebra I strategies" (difference = 0.39, effect size = 0.67, $p < .01$). In addition, a statistically significant higher percentage of treatment teachers (83.7 percent) responded that they would "like to teach a hybrid algebra I course again" than did control teachers (48.8 percent) who would "prefer to use the district's current algebra I curriculum" if they taught algebra I again (difference = 34.9, chi-squared = 11.87, $p < .01$).

Table 4.7. Teacher survey results

Combined cohorts (Treatment n = 45 surveys; control n = 47 surveys)

	Combined cohorts Treatment		Control				
	Mean	**Standard deviation**	**Mean**	**Standard deviation**	**Mean score difference**	**Effect size**	***p***
Dimension							
Impact on instruction (items 2, 3, 9, 11, 15)	3.84	0.58	3.67	0.63	0.17	0.28	0.17
Readiness to implement algebra I (items 7, 8, 10, 12, 13)	4.26	0.53	4.33	0.48	–0.08	– 0.15	0.46
Impact on students (items 1, 4, 5, 6, 16)	3.48	0.62	3.71	0.74	–0.23*	–0.34	0.04
Algebra I activity items							
Teacher use of algebra I strategies (items 17, 18, 19)	4.02	0.46	3.89	0.59	0.13	0.25	0.29
Student use of algebra I strategies (items 20–27)	3.28	0.61	2.90	0.55	0.39*	0.67	< 0.01
Use again?	**Treatment**		**Control**				
	Number	**Percent**	**Number**	**Percent**	**Difference**	**χ^2**	***p***
Would you like to teach a hybrid algebra I/current district algebra I curriculum course again? Scale: 0 = No, 1 = Yes	36	83.7	22	48.8	34.9*	11.87	< 0.01

* Mean score difference (from Wilcoxon-Mann-Whitney test on “dimensions” and “algebra I activity items,” and from chi-square test for “use again?”) is statistically significant at the 95 percent confidence level using a two-tailed test.

Note: Valid n for Hybrid Teacher Questionnaire = 43 treatment teachers, 45 surveys and Control Teacher Questionnaire = 47 control teachers and surveys. Scale: 1 = never, 2 = rarely, 3 = occasionally, 4 = frequently, 5 = extensively. Effect sizes are calculated using Cohen’s d with the pooled standard deviation.

Source: Authors’ calculations based on the data from the Hybrid Teacher Questionnaire and the Control Teacher Questionnaire (2007/08 and 2008/09).

Classroom observation results

External site researchers conducted 165 classroom observations (80 in treatment schools and 85 in control schools). The external site researchers were not informed of the treatment status of schools in which they conducted observations. The School Observation Measure and Algebra I Quality Assessment were used to record data. Up to six School Observation Measure/Algebra I Quality Assessment observations were conducted during one-day visits to each treatment and control school, with a goal of observing all algebra I teachers at least once. Not all teachers in the sample were able to be observed because some were not present on observation day, school administrators did not provide observers with complete rosters of all algebra I teachers, and schools that dropped out of the study did not allow observations. The number of observed teachers was 38 of 63 (60.3 percent) treatment teachers and 44 of 67 (65.7 percent) control teachers, a difference of 5.4 percentage points. Thus, differences across the treatment and control groups on the School Observation Measure and Algebra I Quality Assessment measures should be interpreted with caution.

The 38 treatment teachers and 44 control teachers were observed from one to six times during the one-day visits to their school. Among the 38 treatment teachers, 14 (36.8 percent) had one class observed, 13 (34.2 percent) had two classes observed, and 11 (28.9 percent) had three to six classes observed. Among the 44 control teachers, 21 (47.7 percent) had one class observed, 13 (29.5 percent) had two classes observed, and 10 (22.7 percent) had three to six classes observed.

Summaries of the significant differences between the treatment and control groups are presented below by observation instrument. Complete tables of findings by cohort are in appendix G. Cohen's d effect size is provided as an indication of the magnitude of the difference between the treatment and control groups on each item on the observation instrument and was computed as the mean difference (treatment – control) divided by the pooled standard deviation.

School Observation Measure results

The Wilcoxon-Mann-Whitney test was used to determine if there were statistically significant differences between treatment and control groups on the frequency with which the 11 instructional strategies on the School Observation measure were observed. The instrument used a five-point Likert-type scale where the frequency of activities was rated as 0 = not observed, 1 = rarely, 2 = occasionally, 3 = frequently, and 4 = extensively. As shown in table 4.8A, the analysis revealed statistically significant differences favoring the treatment group for 5 of the 11 items. Treatment teachers were significantly less likely than control teachers to be observed using direct instruction (difference = –0.54, effect size = –0.48, $p = .03$) or independent seatwork (difference = – 1.24, effect size = –1.04, $p < .01$). Treatment teachers were significantly more likely to be observed employing higher-level instructional feedback (written or verbal) to enhance student learning (difference = 0.48, effect size = 0.39, $p = .01$), using the computer for instructional delivery (computer-assisted instruction, drill and practice; difference = 0.99, effect size = 0.59, $p < .01$), and maintaining a high level of student attention/interest/engagement (difference = 0.31, effect size = 0.33, $p = .05$). There were no statistically significant differences between the treatment and control teachers on observations of cooperative/collaborative learning ($p = .54$), use of higher-level questioning strategies ($p = .19$), teacher acting as a coach/facilitator ($p = .78$), student discussions ($p = .06$), technology as a learning tool or resource ($p = .19$), or high academically focused class time ($p = .27$).

Algebra I Quality Assessment results

The Algebra I Quality Assessment was used to collect two types of data: frequency with which the 11 recommended algebra I instructional strategies were used during classroom observations and quality of observed strategies. Each Algebra I Quality Assessment item includes a two-part rating scale: the first indicates whether an activity was observed (0 = no, 1 = yes); the second uses a three-level indicator of quality (1 = low, 2 = moderate, 3 = high) to rate observed strategies. If a strategy was not observed, it was rated "0 = no" and the quality rating for the strategy was left blank. One Algebra I Quality Assessment was completed as part of the observation in each class. Chi-square tests were conducted to determine if there were statistically significant differences in the percent of teachers in the treatment and control groups observed using each of the 11 algebra I activities (three teacher; eight student), and the Wilcoxon-Mann-Whitney test was used to determine if there were statistically significant differences in the quality with which the observed algebra I activities were implemented (low, moderate, high).

There were no statistically significant differences between observations of the treatment and control groups for the three teacher strategies: ask *why* and *what if* questions ($p = .13$), use of number lines, graphs, or diagrams to explain algebra ($p = .98$);

and use of a computer to explain algebra (*p* = .72; table 4.8B). Among the observed student activities, there were no statistically significant differences between the treatment and control groups for five of the eight items: write to explain algebra *(p* = .29); talk to explain algebra (*p* = .71); use things like algebra tiles or blocks (*p* = .59); use graphing calculators (*p* = .27); and use exit slips (*p* = .64). The control group had significantly more instances of student work in groups (difference = –14.7 percent, chi-square = 4.13, *p* = .04) than did the treatment group, whereas treatment classrooms had significantly more observed instances of students using algebra I activities such as “guess and check,” estimating, or drawing (difference = 20.1 percent, chi-square = 8.87, *p* < .01) and students using computers to learn algebra I (difference = 40.0 percent, chi-square = 31.34, *p* < .01) than did control classrooms. An independent *t*-test conducted to determine if the mean numbers of observed activities (teachers or students) were significantly different did not reveal any differences between the treatment and control groups.

Lastly, the quality of implementation of observed strategies was compared for treatment and control groups. Quality was rated on a three-point Likert-type scale with 1 = low, 2 =moderate, and 3 = high for teachers who were observed doing each activity. No quality ratings were given for teachers who were not observed doing each activity. There were no statistically significant differences by treatment status for two of the three teacher activities: ask *why* and *what if* questions (*p* = .97), and use number lines, graphs, or diagrams to explain algebra (*p* = .57; table 4.8C). However, the quality of observed control teacher use of computers to explain algebra was significantly higher than that of treatment teachers (difference = –0.60, effect size = –0.78, *p* < .01). Analysis of the quality of observed student activities revealed no significant differences for five of the eight activities: work in groups (*p* = .65), write to explain algebra (*p* = .11), talk to explain algebra (*p* = .23), use things like algebra tiles or blocks *(p* = .28), and use exit slips (*p* = .20). In contrast, the treatment group was observed with higher quality of implementation for use of activities such as guess and check, estimating, or drawing (difference = 0.98, effect size = 1.55, *p* < .01); use of graphing calculators (difference = 0.48, effect size = 0.75, *p* = .01); and use of computers to learn algebra (difference = 0.50, effect size = 0.95, *p* = .03).

Table 4.8. Classroom observation findings

A. School Observation Measure results

Item		Combined cohorts: Treatment (n = 80) Mean	Treatment Standard deviation	Control (n = 85) Mean	Control Standard deviation	Mean difference	Effect size	p
1	Direct instruction (lecture)	2.53	1.39	3.07	0.84	–0.54*	–0.48	0.03
2	Cooperative/collaborative learning	0.63	1.10	0.78	1.26	–0.15	–0.13	0.54
3	Higher-level instructional feedback (written or verbal) to enhance student learning	1.28	1.29	0.80	1.21	0.48*	0.39	0.01
4	Use of higher-level questioning strategies	1.41	1.25	1.18	1.26	0.23	0.18	0.19
5	Teacher acting as a coach/facilitator	2.21	1.52	2.34	1.30	–0.13	–0.09	0.78
6	Independent seatwork (self-paced worksheets, individual assignments)	0.98	1.27	2.22	1.14	–1.24*	–1.04	< 0.01
7	Student discussion	0.30	0.79	0.12	0.47	0.18	0.28	0.06
8	Computer for instructional delivery (computer-assisted instruction, drill and practice)	2.26	1.76	1.27	1.59	0.99*	0.59	< 0.01
9	Technology as a learning tool or resource (Internet research, spreadsheet or database creation, multimedia, CD-ROM, Laserdisc)	0.75	1.39	0.45	1.07	0.30	0.24	0.19
10	High academically focused class time	3.39	0.74	3.42	0.93	–0.03	–0.04	0.27
11	High level of student attention/interest/engagement	3.09	0.84	2.78	1.02	0.31*	0.33	0.05

* Mean score difference (from Wilcoxon-Mann-Whitney test) is statistically significant at the 95 percent confidence level using a two-tailed test.
Note: Scale: 0 = not observed, 1 = rarely, 2 = occasionally, 3 = frequently, 4 = extensively. Effect sizes are calculated using Cohen's d with the pooled standard deviation.
Source: Authors' analysis of the data from the School Observation Measure (2007/08 and 2008/09).

B. Algebra I Quality Assessment results: observed activities

Item		Combined cohorts: Treatment (n = 80) Number	Treatment Percent	Control (n = 85) Number	Control Percent	Difference (percent)	χ^2	p
Teacher activities								
1	Ask "Why?" and "What if?" questions	56	70.0	50	58.8	11.2	2.24	0.13
2	Use number lines, graphs, or diagrams to explain algebra	51	63.8	54	63.5	0.3	<0.01	0.98
3	Use a computer to explain algebra	28	35.0	32	37.7	–2.7	0.13	0.72
Student activities								
4	Work in groups	19	24.1	33	38.8	–14.7*	4.13	0.04
5	Write to explain algebra (descriptions, poetry, songs, reflections)	4	5.1	8	9.4	–4.3	1.14	0.29
6	Talk to explain algebra	32	40.5	32	37.7	2.8	0.14	0.71
7	Use things like algebra tiles or blocks	4	5.1	6	7.1	–2.0	0.29	0.59
8	Use activities such as "guess and check," estimating, or drawing	28	35.4	13	15.3	20.1*	8.87	< 0.01
9	Use graphing calculators	29	36.7	38	44.7	–80.0	1.22	0.27
10	Use computers to learn algebra	40	50.6	9	10.6	40.0*	31.34	< 0.01
11	Use "exit slips"	5	6.3	7	8.2	–1.9	0.22	0.64

* Difference (from chi-square test) is statistically significant at the 95 percent confidence level using a two-tailed test.
Note: Scale: 0=No, 1=Yes.
Source: Authors' analysis of the data from the Algebra I Quality Assessment (2007/08 and 2008/09).

C. Algebra I Quality Assessment results: quality of observed activities

Item		Combined cohorts								
		Treatment (n = 80)			Control (n = 85)			Effect size	Mean difference	p
		Number	Mean	Standard deviation	Number	Mean	Standard deviation			
Teacher activities										
1	Ask “Why?” and “What if?” questions	56	2.22	0.69	50	2.19	0.82	0.04	0.03	0.97
2	Use number lines, graphs, or diagrams to explain algebra	51	2.24	0.62	54	2.30	0.69	–0.09	–0.06	0.57
3	Use a computer to explain algebra	28	1.93	0.84	32	2.53	0.72	–0.78	–0.60*	< 0.01
Student activities										
4	Work in groups	19	2.37	0.76	33	2.31	0.64	0.09	0.06	0.65
5	Write to explain algebra (descriptions, poetry, songs, reflections)	4	2.50	1.00	8	1.75	0.46	1.23	0.75	0.11
6	Talk to explain algebra	32	1.94	0.72	32	2.16	0.72	–0.31	–0.22	0.22
7	Use things like algebra tiles or blocks	4	1.60	0.89	6	2.17	0.75	–0.77	–0.57	0.28
8	Use activities such as “guess and check,” estimating, or drawing	28	2.29	0.71	13	1.31	0.48	1.55	0.98*	< 0.01
9	Use graphing calculators	29	2.64	0.49	38	2.16	0.75	0.75	0.48*	< 0.01
10	Use computers to learn algebra	40	2.83	0.44	9	2.33	0.87	0.95	0.50*	0.03
11	Use “exit slips”	5	2.14	0.90	7	2.71	0.49	–0.82	–0.57	0.20

* Mean difference (from Wilcoxon-Mann-Whitney test) is statistically significant at the 95 percent confidence level using a two-tailed test.

Note: Scale: 1 = low, 2 = moderate, 3 = high. Effect sizes are calculated using Cohen’s d with the pooled standard deviation.

Source: Authors’ analysis of the data from the Algebra I Quality Assessment (2007/08 and 2008/09).

5. Estimated Impacts of the Intervention

This study examined the effects of the Kentucky Virtual Schools' hybrid program for algebra I on students' math outcomes and grade 10 math course enrollment. This chapter begins by comparing the baseline characteristics of the analysis samples for the treatment and control groups. Next, it presents the impact findings for both outcomes, based on multilevel models that account for the clustering of students within schools. Lastly, it discusses the results of sensitivity analyses conducted to establish the robustness of the impact estimates.

For each outcome, the results are in an impact table. The table shows the regression-adjusted mean between the treatment and control group, and the difference in those means for each outcome. The grade 10 math enrollment outcome includes the odds ratio for the impact estimate. Also included in the tables is a *p*-value, used to assess the statistical significance of the impact estimate. The *p*-value shows how likely it is that one will observe the data if the null hypothesis is true (i.e., there is no treatment effect). An impact is considered statistically significant if it is within a 95 percent confidence interval that does not include 0 using a two-tailed test ($p < .05$).

Baseline characteristics of the analysis samples

The student impact analysis was conducted on students who had valid measures of the outcome variables. Due to missing data, this sample differs slightly from the sample enrolled at baseline. Nonetheless, due to the relatively low missing data rates, the characteristics of the analysis sample (table 5.1) generally mirror those for the original sample enrolled. Importantly, the only statistically significant differences between the treatment and control samples are for gender and one of the pretest measures (EXPLORE). There is a significantly higher proportion of boys than girls in the analysis samples for both math achievement and grade 10 course-taking. However, the difference in the EXPLORE pretest measure is only for the grade 10 course enrollment analysis sample.

Table 5.1. Baseline characteristics of the treatment and control groups for the math achievement and grade 10 math course enrollment analysis samples (percent, unless otherwise noted)

	Math achievement analysis sample				Grade 10 math course enrollment analysis sample			
	Treatment	**Control**	**Difference***	***p***	**Treatment**	**Control**	Difference*	***p***
Student covariates								
Underserved minority	3.98	7.91	–3.92	0.06	4.19	7.89	–3.70	0.08
Male	52.17	48.82	3.34*	0.01	53.16	50.44	2.72*	0.02
Age (mean)	15.41	15.34	0.07	0.46	15.45	15.38	0.07	0.32
Recipient of free or reduced-price lunch	62.35	57.78	4.57	0.09	62.91	59.73	3.19	0.09
Enrolled in Individualized Education Plan	10.16	9.82	0.35	0.92	10.26	10.83	–0.58	0.71
Course level: honors	7.49	13.47	–5.98	0.37	6.82	12.62	–5.80	0.34
KCCT pretest deviation	1.53	1.32	0.21	0.51	0.36	0.20	0.16	0.82
EXPLORE pretest deviation	0.20	0.17	0.03	0.20	0.07	0.02	0.04*	0.03
Sample size	2,847	3,017			3,142	3,267		
School covariates								
School-level KCCT pretest score (mean)	836.99	837.59	–0.60	0.84	836.93	837.31	–0.39	0.85
School-level EXPLORE pretest score (mean)	13.86	13.84	0.02	0.93	13.87	13.81	0.06	0.96
Rural	55.18	68.28	–13.10	0.55	53.98	68.50	–14.52	0.50
Strata 1 (cohort I schools)	49.77	58.60	–8.83	1.00	49.77	58.60	–8.83	1.00
Strata 2 (rerandomized cohort I schools)	7.06	12.63	–5.57	0.75	7.06	12.63	–5.57	0.75
Strata 3 (new cohort II schools)	43.17	28.77	14.40	0.48	43.17	28.77	14.40	0.52
Sample size	24	23			24	23		

* Statistically significant at the 95 percent confidence level.
Source: Authors' calculations based on Kentucky Department of Education data for enrollment (2007/08, 2008/09, and 2009/10) and demographic characteristics (2007/08 and 2008/09), Kentucky Department of Education Commonwealth Academy Commonwealth Accountability Testing System results for KCCT (2006/07 and 2007/08), EXPLORE (2006/07 and 2007/08), and PLAN (2008/09 and 2009/10); and data provided directly by schools (2007/08 and 2008/09).

Confirmatory analysis of impacts

As chapter 3 discussed, two-level hierarchical linear models (for math achievement) or generalized models (for 10^{th} grade math course enrollment) with covariates for both student- and school-level covariates were used for the confirmatory impact estimates. The analysis samples include all treatment and control group students who had data for the respective outcomes and, thus, the reported findings are "intent-to-treat" estimates.

The analyses indicate that the hybrid algebra I intervention did not have a statistically significant impact on either student math achievement or course enrollment. The mean scale scores on the pre-algebra/algebra portion of the PLAN were 6.83 for the treatment

group and 7.09 for the control group. The estimated intent-to-treat impact is –0.25 scale score points (standard error = 0.24). This difference, which translates into an effect size of -0.09, has a p-value of .16, which is well below the critical value of .05 for establishing statistical significance.

Table 5.2. Confirmatory impact estimates

Outcome measure	Treatment group mean (standard deviation)	Control group mean (standard deviation)	Difference[a*] (standard error)	*Odds ratio (standard error)*	p-value	95 percent confidence interval	Estimated impact in effect size units[b]
Math achievement: Grade10 pre-algebra/algebra PLAN	6.83 (2.68)	7.09 (2.91)	–0.25 (0.18)	na	0.16	–0.61, 0.10	-0.15
Grade 10 math course enrollment	0.86	0.86	0.00	1.35 (0.44)	0.36	na	na

a. Impact estimates are estimated using the HGLM regression equation to calculate each student's probability of enrolling in a higher-level math course using individual student characteristics, and then generating group-level means for the treatment and control groups using the treatment variable. The impact estimate is calculated by the first difference between the treatment and control groups.
b. Effect sizes are calculated using Cohen's d with the pooled standard deviation.
* Coefficient (of estimated intent-to-treat impact) is statistically significant at the 95 percent confidence level using a two-tailed test.
Note: Valid *n* for PLAN = 5,863.
Source: Authors' calculations based on Kentucky Department of Education data for enrollment (2007/08 and 2008/09) and demographic characteristics (2007/08 and 2008/09), Kentucky Department of Education Commonwealth Accountability Testing System results for KCCT (2006/07 and 2007/08), EXPLORE (2006/07 and 2007/08), and PLAN (2008/09 and 2009/10); and data provided directly by schools (2007/08 and 2008/09).

An estimated 86 percent of both the treatment and control group students enrolled in Algebra 1 in fall of 10th grade. The adjusted odds ratio was 1.35 to 1 (standard error = 0.44) that treatment students will enroll in a higher grade 10 math course. For both the unadjusted and adjusted impact estimates, there was no statistically significant difference at the 95 percent confidence level with the statistical tests based on standard errors that take clustering into account.

Sensitivity analysis

As discussed in chapter 3, we conducted four sets of sensitivity analyses to establish whether the main impact findings were robust to analytic decisions. Specifically, we examine sensitivity of the findings to the choice of pretest measures included in the model; to the selection of covariates used; to decisions about which students are included in the study sample; and to decisions about which schools are included in the analysis. None of the models produced a change from the results of the confirmatory impact analysis. See tables H1 and H2 in appendix H for the results of the sensitivity analysis.

Models excluding all student- and school-level covariates, except for the student-level pretest scores were estimated first, helping determine whether the impact findings were

sensitive to the model specification. These analyses included three regressions each for the PLAN and grade 10 math course enrollment outcomes; one with both the KCCT and EXPLORE pretests; one with KCCT only; and one with EXPLORE only. The impact estimates were invariant to the exclusion of student- and school-level covariates other than the student-level pretest scores.

The full models including the student- and school-level covariates were estimated second, but with only one set of pretest variables at a time, to determine whether the estimates were sensitive to the type and number of pretest variables included as baseline covariates in the full models. These analyses included two regressions each for the PLAN and grade 10 math course enrollment outcomes; one specification excluding the EXPLORE pretest; and one excluding the KCCT pretest. The impact estimates were insensitive to the type and number of pretest variables included in the estimation.

Estimated next were the full models for both the PLAN and grade 10 math course enrollment outcomes, excluding students not enrolled in a full-year algebra I course that began and ended within the intervention year. In the block (part-year) courses, students were exposed to the intervention for a less time, and teachers in the fall semesters had less experience using the materials and less participation in school-year professional development sessions. This analysis was used to determine if including students from block courses reduced the overall impact of the intervention. The impact estimates were insensitive to excluding students not enrolled in a full-year course.

Finally, the full models for both the PLAN and grade 10 math course enrollment outcomes were estimated, excluding from cohort II the six duplicate schools in the control group in cohort I and then rerandomized in cohort II (Strata 2). One analytic concern was that the randomized observations were no longer independent, because of the presence of these duplicate schools, which might affect the estimated standard errors.[31] Based on the results of these models, the impact estimates and the standard errors were invariant to including in the sample the duplicate schools from strata 2.

[31] Nine students in the rerandomized schools repeated grade 9 and were thus included in the analysis samples for both cohort I and cohort II.

6. Exploratory Impact Analyses

The potential for differential impacts on subgroups for both student achievement outcome measures were explored, including the performance on the pre-algebra/algebra PLAN in the fall of grade 10 and grade 10 math course enrollment. This chapter presents the impact findings for both outcomes and discusses the results from the sensitivity analyses.

For the exploratory research questions, interaction terms were added (one at a time) to the hierarchical linear models used in the confirmatory analysis between the treatment variable and specific subgroup covariates of interest. The interpretation of the findings depends on the statistical significance and direction of the results for both the main effects and the interaction term. A 95 percent confidence interval using a two-tailed test was used to test for statistical significance. As reported in the confirmatory impact estimates, there was no statistically significant main effect of the treatment in the overall sample for either performance on the pre-algebra/algebra PLAN in the fall of grade 10 or math course enrollment in grade 10. While the outcomes of the treatment group as a whole, on average, were not significantly different from those of the control group as a whole, the interaction terms indicated whether the impact of the treatment differs for subgroups of students or schools. For all the student and school subgroups, no statistically significant difference in the impacts of the treatment on math achievement in pre-algebra/algebra on the PLAN or on grade 10 math course enrollment was found.

Estimated impacts for subgroups defined by gender

We estimated impacts for subgroups defined by gender because of the extensive literature suggesting that males and females tend to have different learning styles and different approaches to math problem solving (Carr and Davis 2001; Friedman 1995; Geary et al. 1999). Thus, we hypothesized that males and females might respond differently to the Kentucky Virtual Schools' hybrid program for algebra I. The exploratory analysis indicates, however, that the effects of the hybrid algebra program were not significantly different for boys than for girls (table 6.1). Consistent with the overall impact estimate reported in table 5.2, the estimates for both boys and girls are negative, not statistically significant, and not significantly different from one another.

Table 6.1. Exploratory analysis of impacts on math achievement results, by gender

Student subgroup	Treatment group mean (standard deviation)	Control group mean (standard deviation)	Difference* (standard error)	Estimated impact in effect-size units[a]	*p*
Male	6.82 (2.87)	7.01 (3.06)	–0.19	–0.06	0.03
Female	6.52 (2.47)	6.84 (2.76)	–0.31	–0.12	n/a
Difference	0.30 (2.68)	0.17 (2.91)	0.12 (0.11)	0.04	0.27

a. Effect sizes are calculated using Cohen's d with the pooled standard deviation.
* Difference is statistically significant at the 95 percent confidence level using a two-tailed test.
Note: Means are regression-adjusted using the same set of baseline covariates as the confirmatory analysis. Valid *n* for records for PLAN with nonmissing gender = 5,851 (male = 2,951, female = 2,900). The *p*-value is for the coefficient on the interaction term between treatment and male.
Source: Authors' analysis based on Kentucky Department of Education data for enrollment (2007/08 and 2008/09) and demographic characteristics (2007/08 and 2008/09), Kentucky Department of Education Commonwealth Accountability Testing System results for KCCT (2006/07 and 2007/08), EXPLORE (2006/07 and 2007/08), and PLAN (2008/09 and 2009/10); and data provided directly by schools (2007/08 and 2008/09).

Similarly, the estimated impacts on math course enrollment are not significantly different between males and females ($p = .27$; table 6.2). For males, the predicted probability of enrolling in a higher-level math course in grade 10 was 0.84 for the treatment group and 0.83 for the control group. For females, it was 0.88 for the treatment group and 0.90 for the control group.

Table 6.2. Exploratory analysis of impacts on grade 10 math course enrollment, by gender

Student subgroup	Treatment group probability	Control group probability	Difference[a*]	Odds ratio (standard error)	Log odds ratio	*p*
Male	0.84	0.83	0.01*	1.07	0.07	<0.01
Female	0.88	0.90	–0.01	0.89	–0.11	na
Difference	–0.05	–0.07	0.02	1.20 (0.20)	0.18	0.27

a. Impact estimates are estimated using the regression equation to calculate each student's probability of enrolling in a higher-level math course using individual student characteristics, and then generating group-level means for the treatment and control groups using the treatment variable. The impact estimate is the difference-in-difference, which is the difference between males in the treatment and control groups minus the difference between females in the treatment and control groups.
* Difference is statistically significant at the 95 percent confidence level using a two-tailed test.
Note: Predicted probabilities are regression-adjusted using the same set of baseline covariates as the confirmatory analysis. Valid *n* for records for grade 10 math enrollment with nonmissing gender = 6,395 (male = 3,311, female = 3,084). The *p*-value is for the coefficient on the interaction term between treatment and male.
Source: Authors' analysis based on Kentucky Department of Education data for enrollment (2007/08, 2008/09, and 2009/10) and demographic characteristics (2007/08 and 2008/09), Kentucky Department of Education Commonwealth Accountability Testing System results for KCCT (2006/07 and 2007/08) and EXPLORE (2006/07 and 2007/08), and data provided directly by schools (2007/08 and 2008/09).

Estimated impacts for subgroups defined by enrollment cohort

As noted in chapter 4, there were implementation issues that might be expected to have resulted in differential effectiveness of the hybrid algebra program for those students in cohort I as compared with those in cohort 2. Most notably, by mistake, students in cohort I were given access to a version of the student courseware that contained errors. In addition, both their teachers and their instruction specialists were new to the hybrid algebra program. Thus, it is reasonable to believe that the impacts of the intervention might have been smaller for cohort 1 than II.

In fact, consistent with the null overall findings reported in chapter 4, there were no statistically significant differences in the estimated effects of the intervention between cohort I and II students. The adjusted mean scale score for students in cohort I was 6.71 for the treatment group and 6.86 for the control group, a difference of –0.15 scale score points (table 6.3). For students in cohort II, it was 6.61 for the treatment group and 6.99 for the control group, a difference of –0.38 scale score points.

Table 6.3. Exploratory analysis of impacts on math achievement by cohort

Student subgroup	Treatment group mean (standard deviation)	Control group mean (standard deviation)	Difference* (standard error)	Estimated impact in effect-size units[a]	*p*
Cohort I	6.71 (2.64)	6.86 (2.71)	–0.15	–0.06	0.61
Cohort II	6.61 (2.72)	6.99 (3.15)	–0.38	–0.13	na
Difference	0.10 (2.68)	–0.13 (2.93)	0.23 (.37)	0.08	0.53

a. Effect sizes are calculated using Cohen’s d with the pooled standard deviation.
* Difference is statistically significant at the 95 percent confidence level using a two-tailed test.
Note: Means are regression-adjusted using the same set of baseline covariates as the confirmatory analysis. Valid *n* for records for PLAN = 5,863 (cohort I = 3,184, cohort II = 2,679). The *p*-value is for the coefficient on the interaction term between treatment and the cohort I dummy variable.
Source: Authors’ analysis based on Kentucky Department of Education data for enrollment (2007/08 and 2008/09) and demographic characteristics (2007/08 and 2008/09), Kentucky Department of Education Commonwealth Accountability Testing System results for KCCT (2006/07 and 2007/08), EXPLORE (2006/07 and 2007/08), and PLAN (2008/09 and 2009/10); and data provided directly by schools (2007/08 and 2008/09).

Similarly, there is no significant difference in the estimated impacts on 10th grade course enrollment between cohorts (p = .95; table 6.4).

Table 6.4. Exploratory analysis of impacts on grade 10 math course enrollment, by cohort

Student subgroup	Treatment group probability	Control group probability	Difference[a*]	Odds ratio (standard error)	Log odds ratio	*p*
Cohort I	0.88	0.86	0.02	1.19	0.18	0.87
Cohort II	0.85	0.86	–0.01	0.92	–0.08	na
Difference	0.03	0.00	0.03	0.96 (0.64)	–0.04	0.95

a. Impact estimates are estimated using the regression equation to calculate each student's probability of enrolling in a higher-level math course using individual student characteristics, and then generating group-level means for the treatment and control groups using the treatment variable. The impact estimate is the difference-in-difference, which is the difference between cohort I students in the treatment and control groups minus the difference between cohort II students in the treatment and control groups.

* Difference is statistically significant at the 95 percent confidence level using a two-tailed test.

Note: Predicted probabilities are regression-adjusted using the same set of baseline covariates as the confirmatory analysis. Valid n for records for grade 10 math enrollment is 6,409 (cohort I = 3,437, cohort II = 2,972). The p-value is for the coefficient on the interaction term between treatment and the cohort I dummy variable.

Source: Authors' analysis based on Kentucky Department of Education data for enrollment (2007/08, 2008/09, and 2009/10) and demographic characteristics (2007/08 and 2008/09), Kentucky Department of Education Commonwealth Accountability Testing System results for KCCT (2006/07 and 2007/08) and EXPLORE (2006/07 and 2007/08), and data provided directly by schools (2007/08 and 2008/09).

Estimated impacts for subgroups defined by rural and nonrural location of schools

We hypothesized that impacts might differ between rural and nonrural schools due to differences in both student backgrounds and the schools themselves. Rural areas are often economically depressed and geographically and socially isolated, making it harder to attract teaching candidates from urban areas (McClure, Redfield, and Hammer 2003) and constraining the pool of qualified teaching candidates.

The analysis shows no statistically significant difference in the estimated impacts of the program between rural and nonrural schools (p = .07; table 6.5). The adjusted mean scale score for the rural subgroup was 6.74 for the treatment group and 6.75 for the control group, a difference of –0.01 scale score points. For the nonrural subgroup, it was 6.59 for the treatment group and 7.33 for the control group, a difference of –0.74 scale score points.

Table 6.5. Exploratory analysis of impacts on math achievement results for students in rural and nonrural schools

School subgroup	Treatment group mean (standard deviation)	Control group mean (standard deviation)	Difference* (standard error)	Estimated impact in effect-size units[a]	*p*
Rural	6.74 (2.56)	6.75 (2.80)	–0.01	0.00	0.04
Nonrural	6.59 (2.82)	7.33 (3.01)	–0.74	–0.25	na
Difference	–0.58 (2.68)	0.16 (2.91)	–0.73 (0.40)	–0.26	0.07

a. Effect sizes are calculated using Cohen's d with the pooled standard deviation.
* Difference is statistically significant at the 95 percent confidence level using a two-tailed test.
Note: Means are regression-adjusted using the same set of baseline covariates as the confirmatory analysis. Valid *n* for PLAN = 5,863 (rural = 3,631, nonrural = 2,232). The *p*-value is for the coefficient on the interaction term between treatment and rural.
Source: Authors' analysis based on Kentucky Department of Education data for enrollment (2007/08 and 2008/09) and demographic characteristics (2007/08 and 2008/09), Kentucky Department of Education Commonwealth Accountability Testing System results for KCCT (2006/07 and 2007/08), EXPLORE (2006/07 and 2007/08), and PLAN (2008/09 and 2009/10); and data provided directly by schools (2007/08 and 2008/09).

There was also no statistically significant difference between rural and nonrural subgroups in the impact of the treatment on the outcome for grade 10 math course enrollment. For the rural subgroup, the predicted probability of enrolling in a higher-level math course in the postintervention year was 0.90 for the treatment group and 0.88 for the control group (table 6.6). For the nonrural subgroup, it was 0.81 for the treatment group and 0.82 for the control group. The difference in the treatment impacts between rural and nonrural subgroups was not statistically significant at $\alpha = .05$ with the statistical tests based on standard errors that take clustering into account.

Table 6.6. Exploratory analysis of impacts on grade 10 math course enrollment for students in rural and nonrural schools

School subgroup	Treatment group probability	Control group probability	Difference[a*]	Odds ratio (standard error)	Log odds ratio	*p*
Rural	0.90	0.88	0.02	1.22	0.20	0.08
Nonrural	0.81	0.82	–0.01	0.91	–0.09	na
Difference	0.09	0.06	0.03	1.49 (1.10)	0.40	0.59

a. Impact estimates are estimated using the regression equation to calculate each student's probability of enrolling in a higher-level math course using individual student characteristics, and then generating group-level means for the treatment and control groups using the treatment variable. The impact estimate is the difference-in-difference, which is the difference between rural students in the treatment and control groups minus the difference between nonrural students in the treatment and control groups.
* Difference is statistically significant at the 95 percent confidence level using a two-tailed test.
Note: For each subgroup, the predicted probabilities are regression-adjusted using the same set of baseline covariates as the confirmatory analysis. Valid *n* for grade 10 math enrollment = 6,409 (rural = 3,934, nonrural = 2,475. The *p*-value is for the coefficient on the interaction term between treatment and rural.
Source: Authors' analysis based on Kentucky Department of Education data for enrollment (2007/08, 2008/09, and 2009/10) and demographic characteristics (2007/08 and 2008/09), Kentucky Department of Education Commonwealth Accountability Testing System results for KCCT (2006/07 and 2007/08) and EXPLORE (2006/07 and 2007/08), and data provided directly by schools (2007/08 and 2008/09).

7. Summary of Findings and Study Limitations

This study was designed to quantify, through a rigorous RCT, the effectiveness of the Kentucky Virtual Schools' hybrid program for algebra I on grade 9 students' math outcomes. This was a well-established program structured on research-based strategies. It included an Internet-based resource for students and ongoing professional development focused on both content and pedagogy for teachers. The program was evaluated using an experimental design with a large sample size. Sample attrition rates were low and evenly balanced between the treatment and control groups. Multilevel models were used to estimate the impact of the program on students' math achievement in pre-algebra/algebra in the fall of grade 10 on the PLAN, as well as on students' grade 10 math course enrollment. The study found no evidence that the program improved either math achievement or the probability of enrolling in a math course above algebra I in 10th grade. Furthermore, there was no evidence of positive (or negative) impacts for student subgroups defined by gender, enrollment cohort, or rurality of the school setting.

There are limitations due to the fidelity of implementation as 4 of the 24 treatment schools were noncompliant, which means they did not participate in the intervention. In addition, 20 of the 63 teachers did not participate in any component of the intervention; 6 because their school withdrew from the study after randomization, and 14 who attended no professional development sessions and did not use the online student courseware. Twelve of those teachers were never reported by participating schools as teaching algebra I, and were not identified by the research team until student enrollment records were gathered from the Kentucky Department of Education late in the school year. Less than 50 percent of the treatment sample (47.6 percent of teachers in the summer and 42.9 percent of teachers in the school-year) had high or moderate attendance for the professional development components. The percent of teachers rated as low in engagement during the professional development sessions was 19 percent in the summer and 30 percent in the academic year. Further, 65 percent of treatment students had ratings of none or low for use of the Kentucky Virtual Schools' online algebra I materials during the period for which use data were collected. The data available for fidelity of implementation included the number of student connections per week from the electronic archives of the courseware, but not the actual amount of class time spent using the courseware so this is only an approximation of whether the online materials were used as intended. The low levels of fidelity of implementation limit the conclusions that can be drawn about the potential effectiveness of both the professional development program and the online student courseware. Treatment conditions were similar to a real-world setting so this level of fidelity of implementation may be expected if the program was offered by the state.

This study has also several limitations relating to the generalizability of the findings to other settings and contexts. One is that the schools in the sample volunteered to participate. These schools may differ from the broader population of Kentucky high schools on both observable and nonobservable characteristics. As a result, the findings are limited to the study schools and are not generalizable beyond the sample.

The student courseware intended for use in this study was the latest version available from the National Repository of Online Courses when the intervention began. However, the virtual school mistakenly installed an early version of the student courseware. The older courseware had not been through careful review, and student exercises had numerous errors. Complaints from teachers to instructional specialists led to an inquiry by the research team, the discovery of the source of the problem, and the problem's resolution. Updated student courseware was installed at the start of the second semester of the first cohort and used for the rest of the study. While one might expect such an error to negatively affect continued use of the intervention, no such evidence was found in the data. Rates of participation in professional development and frequency of student logins to the courseware were similar for both cohorts. There were also no statistically significant differences between the impacts of the intervention by cohort on math achievement in pre-algebra/algebra on the PLAN or on grade 10 math course enrollment.

Other limitations pertain to the students included in the study. Data are missing on the PLAN outcome for students who are not promoted, leave the Kentucky public school system, or miss testing in grade 10. This means that the results from the analysis of student math achievement apply only to students enrolled in algebra I at a participating school in grade 9 and promoted to grade 10. They do not generalize to algebra I students in participating schools not promoted to grade 10. This limitation affects only a small subset of the study sample, as attrition rates are 16 percent or less on each outcome and evenly balanced between the treatment and control groups.

Also, the results apply only to grade 9 students enrolled in courses leading to algebra I credit. The results should not be generalized to algebra I courses that do not fulfill the algebra I requirement, such as algebra I part A (the first course in a two-year sequence) or algebra I lab (an elective course). Further, the results do not apply to online courses for other subjects or grade levels that are provided through Kentucky Virtual Schools.

A final note: As with RCTs in general, this study does not attempt to measure the impact of the intervention on the subset of treatment schools that used the intervention as intended. Instead, it measures the impact of the offer of treatment on outcomes. The findings for this intent-to-treat analysis are suggestive of the impact on outcomes one might observe at the end of the first year of implementation if the Kentucky Department of Education offered the intervention to similar schools in Kentucky.

Appendix A. Power Analysis

A power analysis was conducted before study recruitment to determine an appropriate sample size. This appendix begins by describing the assumptions underlying the power analysis for the overall program effects, illustrates the power curves used to determine adequate statistical power, and compares the expected sample to the actual sample. This is followed by power analyses for subgroup effects included in chapter 6.

Power analysis for overall impact of intervention on student achievement: assumptions

Students at level 1

It was assumed that, on average, each school has approximately 170 algebra I students. This assumption reflected expected enrollments, based on the size of schools that volunteered to participate early in recruiting. In the actual sample, each school had an average of 148 algebra I students.

Intraclass correlation

It was assumed that individual students would account for 90 percent of variation in the outcome variable, and that clustering at the school level would account for the remaining 10 percent. This means that the intraclass correlation was estimated to be 0.10, based on the consideration that most targeted schools were high-needs high schools in rural Kentucky. As such, some uniformity manifests across the targeted schools, so the variation among them should be smaller than the variation among all high schools in one state. Thus, the hypothesized intraclass correlation of 0.10 (10 percent of variation in the data) was believed to be reasonable and unlikely to underestimate the clustering effect.

Cluster-level covariate

The initial plan was to use at least one aggregated second-level covariate in the model. The cluster-level covariate considered was the pretest math scores aggregated at the school level (the average pretest math score for all students in a school). Research on school academic achievement has repeatedly shown that pretest scores are highly correlated with posttest scores and thus substantially correlated with cluster (school) means of the posttest scores. The initial assumption was that this relationship is R^2_{L2} = 0.49.

Level of significance, one- versus two-tailed tests

The plan was to use $\alpha = 0.05$ as the level of significance. All tests would be two-tailed. If a one-tailed test was desired, the power level would be higher than the power curves shown in the graph.

Effect-size magnitude

Two effect-size levels were assumed: $\delta = 0.20$ and $\delta = 0.25$. A study by Bloom, Hill, Black, and Lipsey (2008) on achievement effect-size benchmarks for education interventions found that the average annual gain in effect size for nationally normed math

tests is 0.25 for students transitioning between grades 9 and 10. In this sense, the study is powered to detect an effect size that represents at least an 85 percent improvement over the annual gain otherwise expected for grade 9 math students.

Potential attrition at different levels (student, school)

The plan was for a randomized controlled trial (RCT) in which all eligible courses culminated in the completion of algebra I credit during the intervention year. Most school attrition occurs across school years (students moving, teachers changing jobs, schools changing administrators), so this one-year treatment study avoids many potential pitfalls of attrition. Still, attrition was carefully considered in planning. For this study, potential attrition might occur at two levels: student and school. At the student level, given the very large number of students, the effect of attrition on statistical power was assumed to be inconsequential. Second, at the school level, planning allowed for attrition of up to 20 percent (up to 9 of the 47 schools; 19 percent), leaving an analysis sample of 38 schools. Researchers felt that 20 percent attrition in one school year might be an overestimate, but it remained within What Works Clearinghouse guidelines (U.S. Department of Education n.d.) for attrition from RCTs with results that are considered strong evidence of impact without reservation.[32]

Number of schools

The plan was to randomize 47 schools with approximately 6,000 students, 24 into the treatment condition (hybrid approach) and 23 into the control condition (business-as-usual).[33]

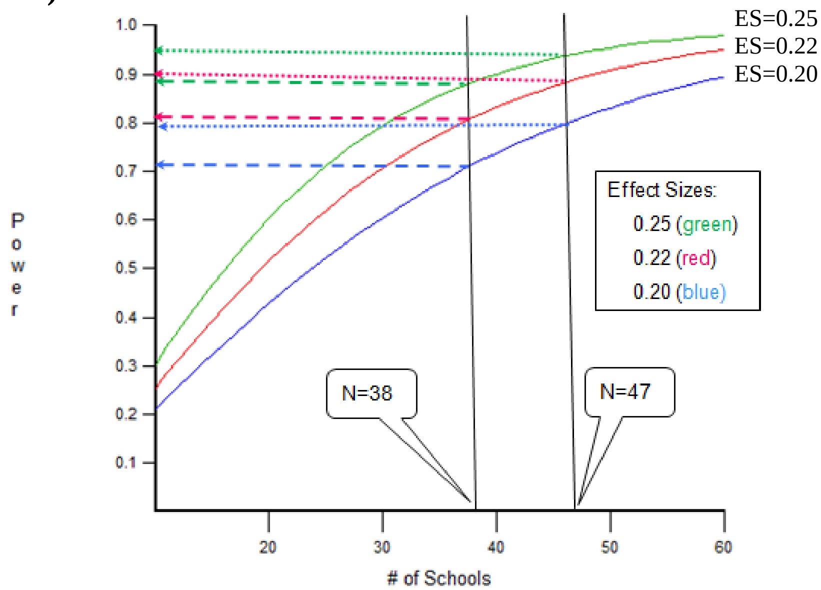
Power curves for cluster randomized trials

The power curves are based on Raudenbush, Spybrook, Liu, and Congdon (2005). Using a two-tailed test of statistical significance, the randomized schools provided adequate statistical power for detecting effect sizes of 0.20 and above (figure A1).

[32] What Works Clearinghouse reviewers assume that bias associated with attrition is minimal when overall attrition is less than 20 percent. See, for example, the review protocol for middle school math interventions at http://ies.ed.gov/ncee/wwc/PDF/msm_protocol_1.1.pdf.

[33] The 47 schools include six duplicate schools. These schools are in the cohort I control group and then included in cohort II and rerandomized. The power analysis does not account for the possible effect of inclusion of these duplicates. (The researchers are unaware of a methodology that would adjust the power estimates based on correlation among of some of the study schools.) The power analysis shows the range of statistical power associated with 0 to 20 percent attrition, which incorporates the statistical power that would pertain if the duplicate schools were deleted from the sample. In addition, the sensitivity tests include the estimation of a model that excludes these duplicate schools.

Figure A1. Power curves using a two-tailed test, overall impact analysis (alpha = 0.05)



Source: Authors' analysis based on Raudenbush, Spybrook, Liu, and Congdon (2005).

Figure A1 shows that, for the lower-bound plan of 38 schools, the power for detecting $d = 0.25$ is 0.88, for detecting $d = 0.22$ is 0.81, and for detecting $d = 0.20$ is 0.72. Thus, the expected sample was adequate for detecting a minimally detectable effect of 0.22 standard deviation.

Actual sample

Because the PLAN is given statewide and outcome data was collected directly from the Kentucky Department of Education, there was no attrition at the school level. Even noncompliant schools could be included in the analysis. For the sample of 47 randomized schools, the power to detect an effect size of 0.20 standard deviation is 0.79, of 0.22 standard deviation is 0.90, and of 0.25 standard deviation is 0.95. Thus, the minimum detectable effect for the randomized sample is about 0.20 standard deviation.

Table A1 summarizes the recruiting sample for all randomized schools and the analysis samples for the primary confirmatory outcome measures.

Table A1. Number of schools in the sample for the confirmatory analyses, by cohort and treatment status

	Sample (schools)			
Full sample	**Cohort**	**Treatment**	**Control**	**Total**
Total number of randomized schools	I: Strata 1 (cohort I schools)	13	12	25
	II: Strata 2 (rerandomized cohort I schools)	3	3	6
	II: Strata 3 (new cohort II schools)	8	8	16
	II: Total	11	11	22
Sample size	Total for 3 strata	24	23	47

Source: Authors' compilation.

Defining effect size for dichotomous variable, grade 10 math course enrollment

One analysis involved a dichotomous variable at the student level (whether students enrolled in a course above algebra I in the postintervention school year). Multilevel logistic regression analysis was used for this dichotomous outcome variable. As discussed in the literature, the effect size of choice for a dichotomous variable is derived from the odds ratio of the treatment condition to the control condition in relation to the dichotomous outcome (enrollment or no enrollment in a math course above algebra I in grade 10). The odds ratio can be transformed into log odds ratio, ln(OR), which would conceptually align with the standardized mean difference. To make the ln(OR) comparable to the standardized mean difference such as Cohen's d, the transformation proposed by Cox (1970) can be applied:

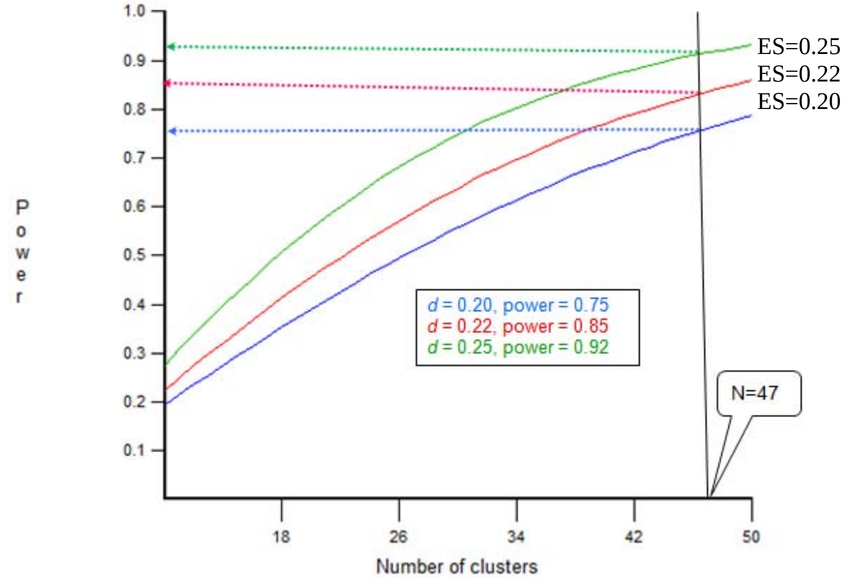
$$LOR_{Cox} = \ln(OR) / 1.65$$

Once this effect-size measure for the dichotomous outcome variable is defined as above, the power analysis proceeds in the same way as for a conventional effect-size measure, such as standardized mean difference between the treatment and control conditions. Hence, the assumptions, power curves, and minimum detectable effects in this appendix apply to both the dichotomous and continuous outcome variables. These assumptions are based on $\alpha = 0.05$ with a two-tailed test.

Power analysis for student subgroups (gender and cohort)

In each school, there were approximately 150 students. It was assumed that 50 percent were male and 50 percent were female; that is, 75 male students and 75 female students in each cluster (school). Also assumed was that both cohorts would have approximately the same number of students. All other assumptions were the same as in the confirmatory power analysis for the overall impact analysis. Figure A2 shows the resulting power curves.

Figure A2. Power curves for student subgroup analysis using a two-tailed test, (alpha = 0.05)



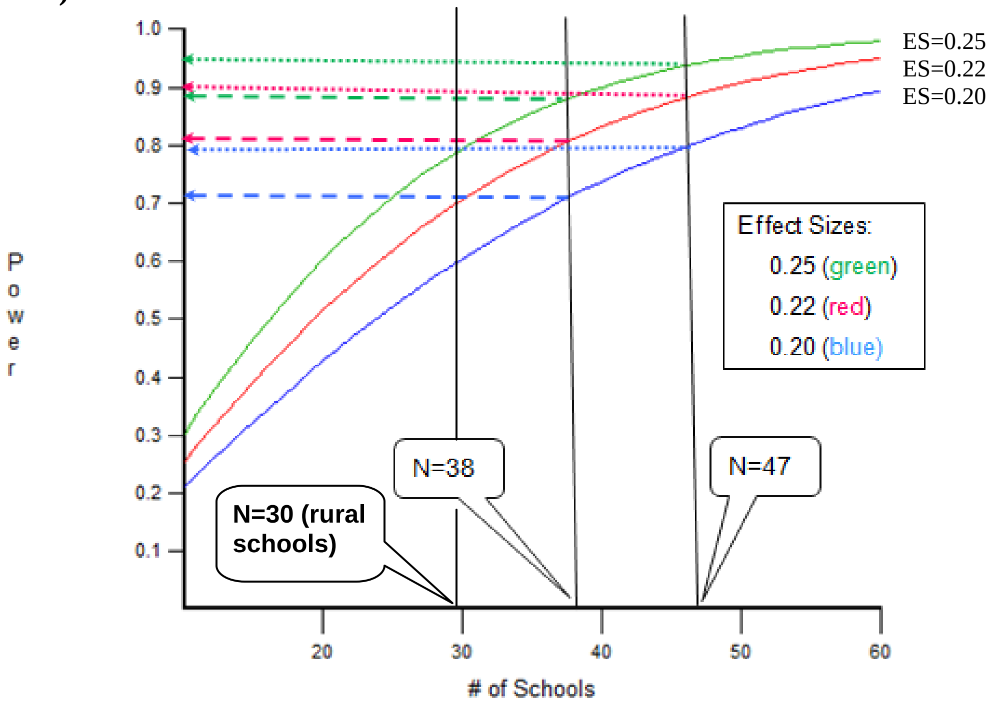
Source: Authors' analysis based on Raudenbush, Spybrook, Liu, and Congdon (2005).

For the sample of 47 schools, the power to detect an effect size of 0.20 standard deviation is 0.75; for an effect size of 0.22 standard deviation, the power is 0.85; for an effect size of 0.25 standard deviation, the power is 0.92. The effect size for this sample associated with 0.80 power is 0.21 standard deviation, a conservative estimate based on the impact of the intervention on a single subgroup.

Power analysis for school subgroups (rural status)

Thirty of the 47 schools in this study were in rural areas. To determine the minimum detectable effect for the subsample of rural schools, the assumptions for the overall analysis were retained, but the number of clusters was reduced from 47 to 30 (figure A3). This is a conservative estimate based on the impact of the intervention on a single subgroup. The minimum detectable effect for rural schools is 0.25 standard deviation.

Figure A3. Power curves for school subgroup analysis using a two-tailed test, (alpha = 0.05)



Source: Authors' analysis based on Raudenbush, Spybrook, Liu, and Congdon (2005).

Appendix B. Data Collected but Not Analyzed

Some data collected for several outcome variables were not analyzed. This appendix discusses these data and the rationale for excluding them from the analyses. All data collected but not analyzed were included in the restricted-use data file.

PLAN composite scores in math, science, and reading

The American College Testing PLAN (PLAN) is used to assess students' knowledge and skills in the fall of grade 10. The Kentucky Department of Education provided the PLAN data with the scale scores for the pre-algebra/algebra strand and the composite scores in math, science, and reading.

Rationale for excluding the PLAN composite scores from the analyses

The initial plan was to examine treatment effects on different parts of the PLAN in the exploratory analysis if statistically significant effects were found in the confirmatory analysis of the PLAN on pre-algebra/algebra outcomes. Researchers speculated that the intervention might have a positive impact on these outcomes because it might develop analytic skills and technical literacy that would enhance student performance on other subjects. After further consideration, it was decided that these outcomes are distal from the intervention's primary intent of increasing student learning and achievement in algebra I. Because of their hypothesized tenuous link to the intervention, the PLAN composite scores were excluded from the exploratory analysis.

Intervention-year assessment of student achievement

Site researchers administered a 50-question "Assessment of Algebraic Understanding" (Educational Testing Service 2004) as an end-of-course assessment in May of grade 9. This assessment used the first half (25 questions) of an end-of-course assessment developed by the Educational Testing Service. The score equals the percentage of questions the student answered correctly.

The Educational Testing Service assessment comprises two psychometrically equivalent halves of 25 questions each. Each half takes 40 minutes to administer, permitting a reliable assessment in a single class period. According to the Educational Testing Service, the reliability of the full assessment is 0.87. Using the Spearman Brown Prophecy formula, the reliability of the assessment when only half the questions are used is 0.76 (Crocker and Algina 1986).

End-of-course assessment administration procedures

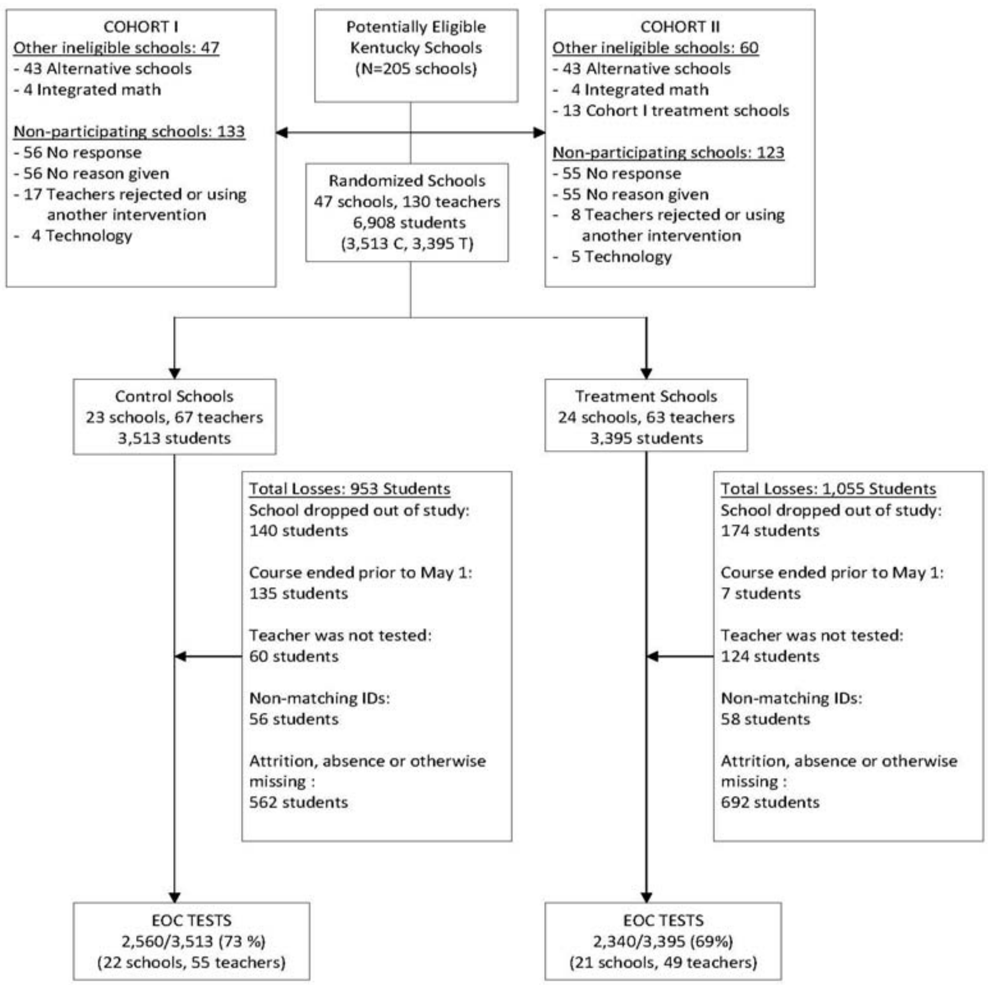
On a prescheduled day(s) at the end of the intervention school year, site researchers administered the algebra I end-of-course assessment in each study class during its regularly scheduled time period. The test schedules were determined from teacher-submitted schedules, and the teachers knew in advance when the test would be given.

Each treatment and control class was assigned a unique seven-digit number identifying the study group, school, teacher, and class section. Treatment and control teachers received an end-of-course assessment packet for each algebra I class section they taught. The packet, labeled with the seven-digit number, contained student coversheets, end-of-course assessment answer sheets, and teacher instructions.

Rationale for excluding the intervention-year end-of-course assessment from the analyses

The Educational Testing Service end-of-course assessment was dropped. Instead, the pre-algebra/algebra portion of the PLAN was used for both the confirmatory and exploratory analyses because the end-of course assessment had a higher nonresponse rate and greater differential attrition by treatment status, which threatened the validity of the randomization. As shown in figure B1, 73 percent of the control group and 69 percent of the treatment group had end-of-course assessment records in the combined cohorts.

Figure B1. Sample detail for end-of-course assessment in grade 9



Source: Authors' calculations based on Kentucky Department of Education data for student enrollment (2007/08 and 2008/09) and researcher collected end-of-course data (2007/08 and 2008/09). Adapted from the Consolidated Standards on Reporting Trials (CONSORT) statement (Altman et al. 2001).

Missing end-of-course assessment data arose from several identifiable sources:

- *Schools were not tested.* In cohort I, four schools dropped out of the study at the beginning of the intervention school year: 314 students (140 control, 174 treatment) were not tested because they were enrolled in a school that withdrew from the study.[34]
- *Courses ending before May in cohort I.* In the first semester of this project, the researchers lacked clearance from the Office of Management and Budget to begin data collection and thus could not enter cohort I schools to conduct end-of-course assessments. (So testing did not occur in the fall of 2007/08.) In cohort I, 142 students (135 control, 7 treatment) were not tested because they were enrolled in courses that ended before May 1. Students were tested in both the fall and spring semesters in cohort II.
- *Teachers were not tested.* Site researchers attempted to test all algebra I classes, but the Kentucky Department of Education enrollment records indicated that not all teachers were tested. There were three reasons why teachers were missing test scores: teachers who did not participate in the intervention did not agree to be tested; the school's list of algebra I teachers was incomplete, so the site researchers were not aware of all teachers who needed to be tested; and courses had fewer than five students, with no students available on the day that site researchers were scheduled to administer the test. Among the students enrolled in a course ending in May at a school that did not drop out of the study, 60 control students and 124 treatment students had a teacher who was not tested.
- *Nonmatching identification numbers on end-of-course assessment records.* Student identification numbers were entered by hand on end-of-course assessment response sheets, a practice that invites errors in data entry. Researchers sought to correct any nonmatching numbers by manually searching in the same school for enrollment records that were missing a record and had an almost identical number. After this data-cleaning process, 114 student identification numbers (56 control, 58 treatment) from the end-of-course assessment records still matched no enrollment, demographic, or pretest records.
- *Student attrition, absence, or otherwise missing.* Some students enrolled in algebra I at the beginning of the school year did not complete the course. Other students may have been enrolled but were absent on the testing day. It is also possible that students were missing records for such reasons as not returning the answer sheet. A total of 1,254 students (562 control, 692 treatment) were in this category of missing data.

Postintervention-year assessment of teacher effectiveness

The same end-of-course assessment administered to grade 9 students at the end of the intervention year was also scheduled for a new cohort of grade 9 students served by all participating teachers at the end of the second school year after the intervention. Treatment schools were given access to the intervention materials in the postintervention

[34] Fewer than three schools in cohort II withdrew, but they still allowed the researchers to collect end-of-course assessment data.

year and were encouraged—but not required—to use them to stay in the study. Even if teachers did not use the material as intended (for example, if they used the materials exclusively for whole-group instruction rather than for individual students), students could benefit from the teacher professional development provided in the first year of the intervention. The postintervention-year end-of-course assessment was expected to be an indicator of whether the intervention had a sustained impact on teacher instructional practices in the second year after exposure to the treatment, whether or not the teacher continued to use the intervention materials.

Rationale for excluding the postintervention-year end-of-course assessment from the analyses

The initial plan was that the exploratory analyses would evaluate the sustained effects of the intervention on teacher instructional effectiveness, measured in terms of student achievement on the grade 9 end-of-course assessment at the end of the postintervention school year. The postintervention-year end-of-course assessment, however, was dropped from the analysis due to the high nonresponse rate and differential attrition by treatment status, which threatened the validity of the randomization.

Data were collected for the postintervention-year end-of-course assessment for cohort I. Because it was decided late to collect data for year 2, the two-year commitment was not emphasized to teachers in this cohort, which might have led to very low compliance rates for cohort I year 2. The postintervention end-of-course assessment collection was also subject to attrition, due to schools dropping out of the study. For the postintervention-year end-of-course assessment, the response rate in cohort I was 53 percent for the treatment group and 74 percent for the control group.

The postintervention-year end-of-course assessment was not administered for cohort II. Weighing against collection of the remaining data were the cost and the need to drop the cohort I data from any analyses because of the high, unbalanced attrition rate. If these outcome data were dropped for cohort I, the remaining sample would be underpowered for conducting a statistical analysis of impact. Timing of the second data collection (May 2010) was also a concern, as the late collection date threatened completion of the study within the contract period.

High school continuation

Kentucky Department of Education data on school enrollment and withdrawal records in year 2 were collected for both cohorts. Students who did not stay in the same school in year 2 showed a withdrawal code indicating that they dropped out of school ("dropout-age 16 or older") or left for a different reason, such as transferring to another school.

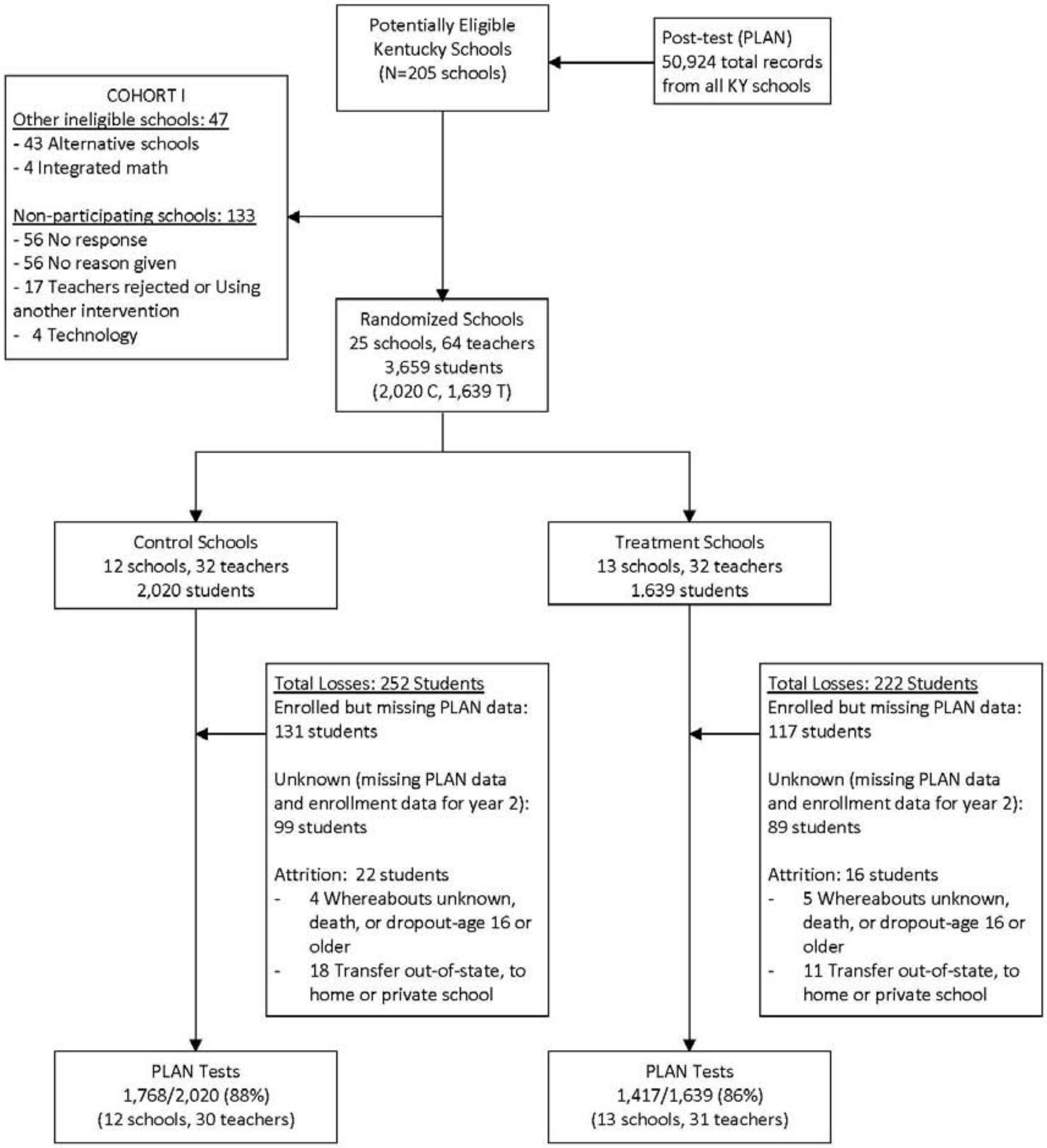
Rationale for excluding high school continuation from the analyses

The initial plan was to include an exploratory research question examining the intervention's effect on students' educational persistence, measured in terms of high school continuation in January of the postintervention school year. The high school continuation outcome, however, was excluded from the exploratory analysis because of the hypothesized tenuous link to the intervention.

Appendix C. Sample Detail

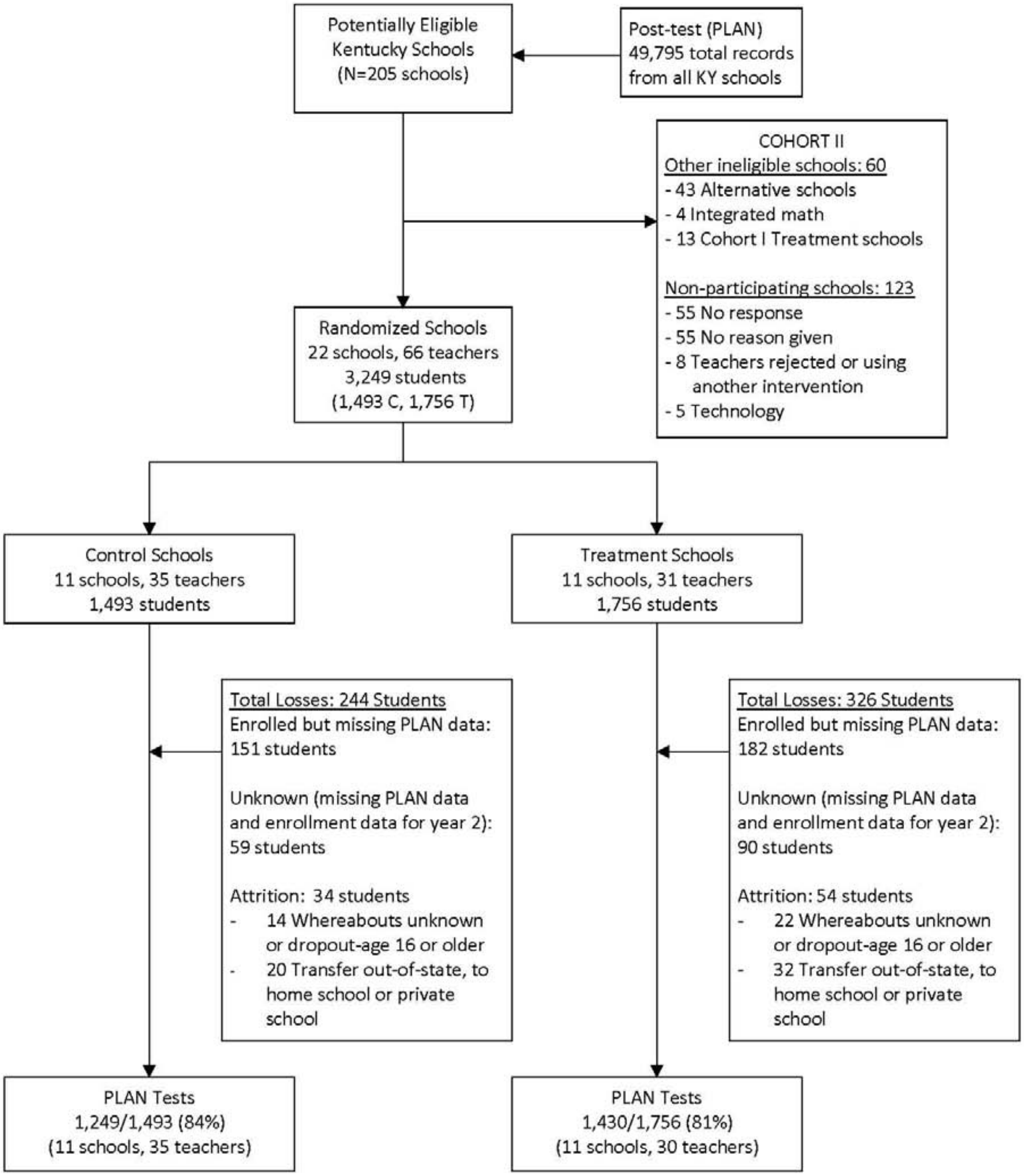
This appendix includes further information on the study sample, separately for each cohort. The assignment of study participants and participant losses is summarized in flowcharts adapted from the Consolidated Standards on Reporting Trials (CONSORT) statement (Altman et al. 2001). Figures C1–C4 show the sample detail for the PLAN math achievement sample and the sample detail for the grade 10 math course enrollment sample, with separate figures for each cohort for each outcome measure. The response rate for cohort I was 88 percent (1,768 / 2,020) for the control group and 86 percent (1,417 / 1,639) for the treatment group. For cohort II, it was 84 percent (1,249 / 1,493) for the control group and 81 percent (1,430 / 1,756) for the treatment group. The enrollment response rates for cohort I control group (1,903 / 2,020) and treatment group (1,534 / 1,639) were both 94 percent. For cohort II, the response rate for the grade 10 math course enrollment was 91 percent (1,364 / 1,493) for the control group and 92 percent (1,608 / 1,756) for the treatment group.

Figure C1. Sample detail for the cohort I pre-algebra/algebra PLAN assessment



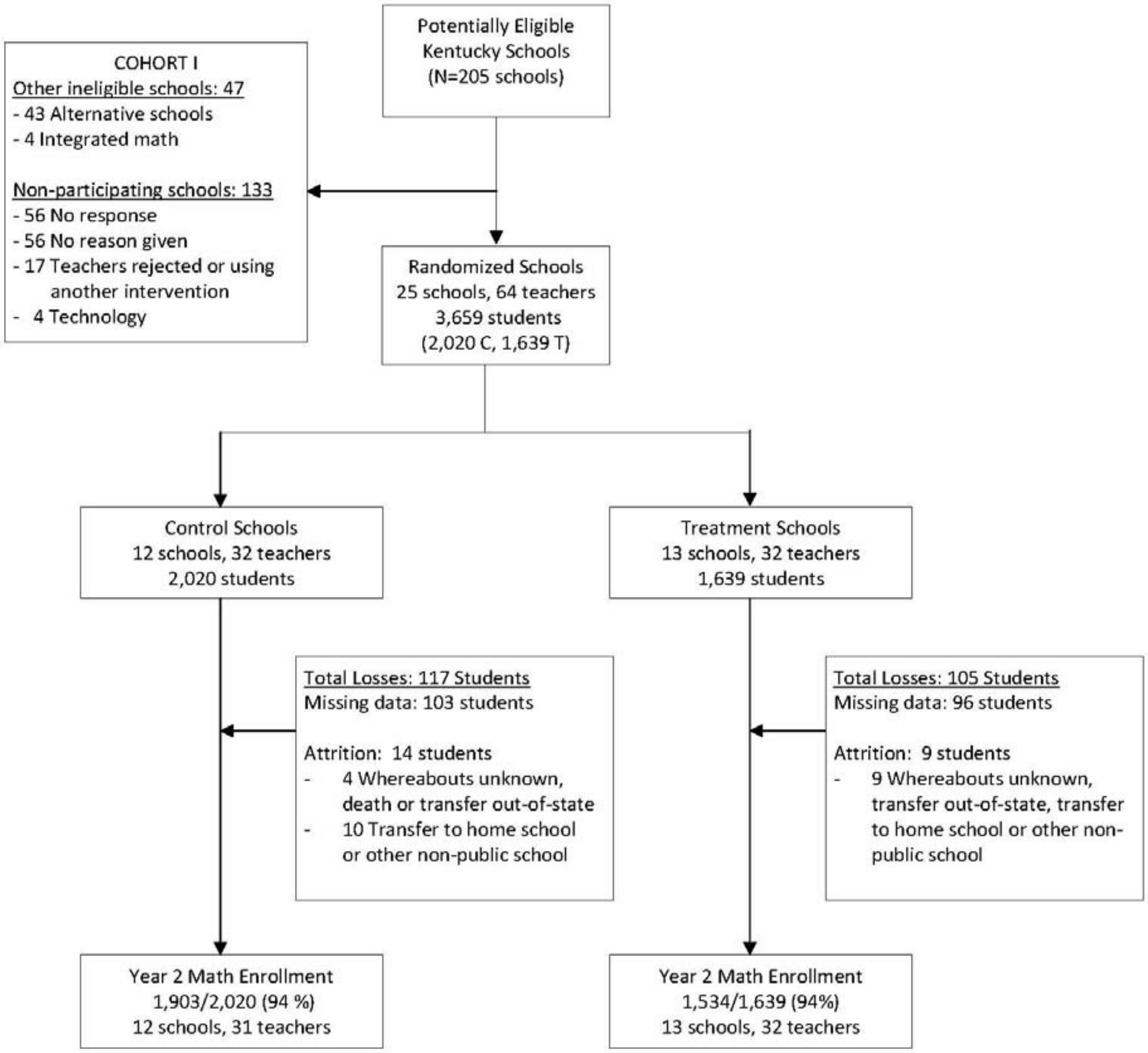
Source: Authors' calculations based on Kentucky Department of Education data for student enrollment (2007/08 and 2008/09) and Kentucky Department of Education Commonwealth Accountability Testing System results for PLAN (2008/09). Adapted from the Consolidated Standards on Reporting Trials (CONSORT) statement (Altman et al. 2001).

Figure C2. Sample detail for the cohort II pre-algebra/algebra PLAN assessment skills



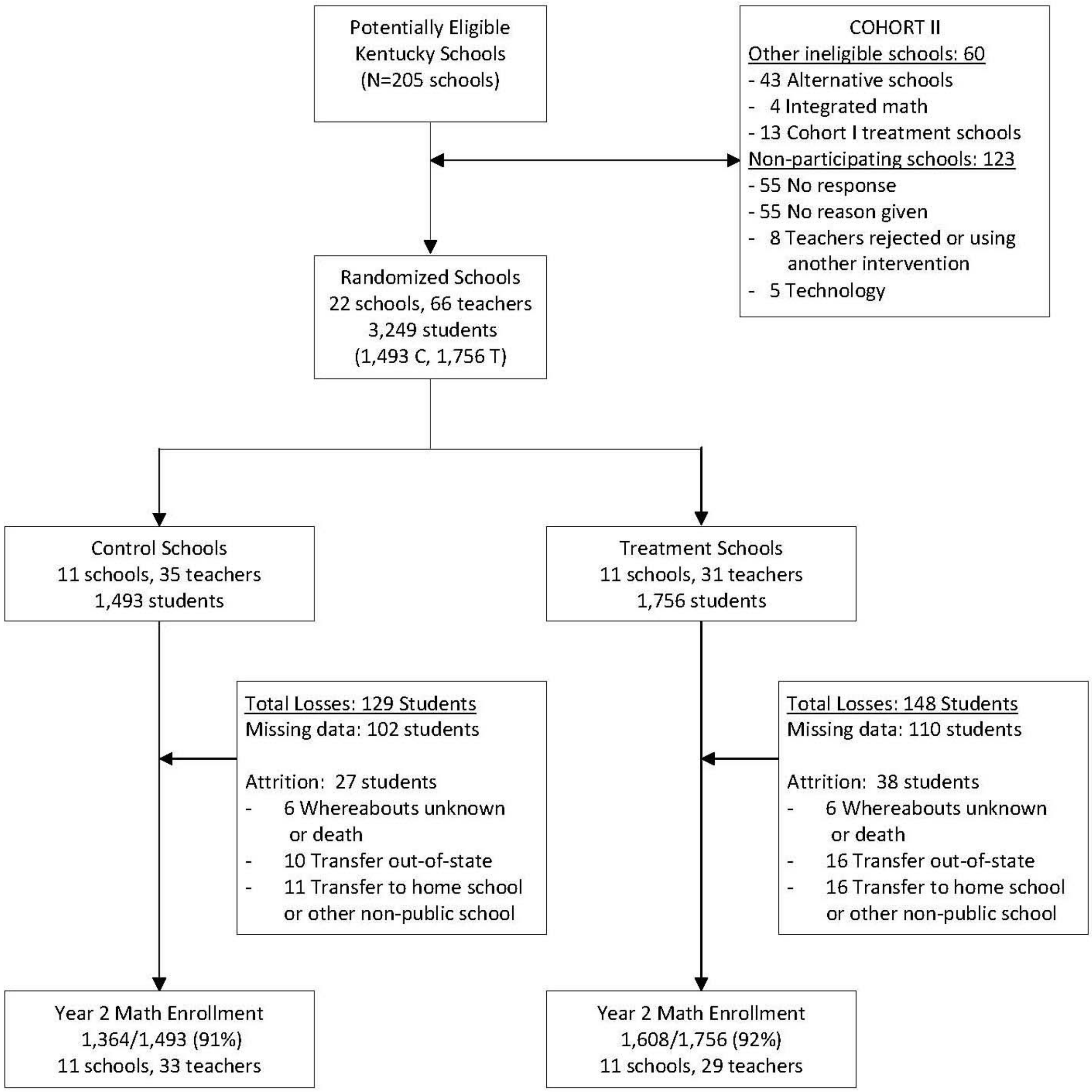
Source: Authors' calculations based on Kentucky Department of Education data for student enrollment (2008/09 and 2009/10) and Kentucky Department of Education Commonwealth Accountability Testing System results for PLAN (2009/10). Adapted from the Consolidated Standards on Reporting Trials (CONSORT) statement (Altman et al. 2001).

Figure C3. Sample detail for the cohort I grade 10 math course enrollment



Source: Authors' calculations based on Kentucky Department of Education data for student enrollment (2007/08 and 2008/09). Adapted from the Consolidated Standards on Reporting Trials (CONSORT) statement (Altman et al. 2001).

Figure C4. Sample detail for the cohort II grade 10 math course enrollment



Source: Authors' calculations based on Kentucky Department of Education data for student enrollment (2008/09 and 2009/10). Adapted from the Consolidated Standards on Reporting Trials (CONSORT) statement (Altman et al. 2001).

Appendix D. Technical Information

This appendix details technical information on the methodology discussion in chapter 3. It describes the reliability of the teacher survey instruments; the confirmatory impact analysis with complete model specifications; and the procedures for handling missing data, conducting sensitivity analyses, making adjustments for multiple comparisons, and weighting.

Reliability of the teacher surveys

Teacher surveys, along with classroom observations (Ross, Smith, Alberg, and Lowther 2004), were used to describe differences in the treatment and control conditions, by collecting teacher perceptions of the algebra I approach they used (hybrid or district curriculum)[35]. The two versions of the instrument, the Hybrid Algebra 1 Teacher Questionnaire (Hybrid Teacher Questionnaire) and the Algebra 1 Control Teacher Questionnaire (Control Teacher Questionnaire), were an adaptation of the 20-item validated Teacher Technology Questionnaire (Inan, Lowther, Ross, and Strahl 2010), designed to assess teacher perceptions of technology integration on five subscales: impact of technology on students, impact of technology on instruction, readiness to integrate technology, technical support, and overall support (Lowther and Ross 2000). The Teacher Technology Questionnaire reliability coefficient was calculated separately for each subscale, indicating high reliability for each subscale of the instrument, ranging from 0.75 to 0.89 (Sterbinsky and Burke 2004).

The reliability of the Hybrid Teacher Questionnaire was calculated using 47 teacher participants; the Control Teacher Questionnaire, 49 participants. All had completed the instruments. The internal consistency reliability coefficient (Cronbach's alpha) was calculated separately for each subscale.

The reliability estimates on the Control Teacher Questionnaire were higher than the Hybrid Teacher Questionnaire on three subscales: impact on students (0.81 compared with 0.63), impact on instruction (0.76 compared with 0.73), and teacher use of strategies (0.47 compared with 0.38); those on the Hybrid Teacher Questionnaire were higher on two subscales: readiness to teach algebra I (0.64 compared with 0.61) and student use of strategies (0.71 compared with 0.65; table D1). Based on the standard reliability cutpoint of 0.70 for Cronbach's alpha, the reliability was low for both questionnaires on teacher use of strategies and readiness to teach algebra I. Reliability was also low for the impact on students subscale on the Hybrid Teacher Questionnaire (0.63).

Appendix G1 provides detailed information regarding the Hybrid and Control Teacher Questionnaire items and results and appendix G2 provides details regarding classroom observation instruments and results.

[35] Any inquiries regarding the teacher surveys or classroom observation instruments and methodology used for this project may be directed to: The Center for Research in Educational Policy, The University of Memphis, 325 Browning Hall, Memphis, TN 38152.

Table D1. Hybrid Teacher Questionnaire and Control Teacher Questionnaire reliability analysis

Scale	Number of items	Hybrid Teacher Questionnaire		Control Teacher Questionnaire	
		n	**Cronbach's Alpha**	***n***	**Cronbach's Alpha**
Impact on students	5	44	0.63	49	0.81
Impact on instruction	5	44	0.73	47	0.76
Readiness to teach algebra I	5	47	0.64	47	0.61
Teacher use of strategies	3	47	0.38	49	0.47
Student use of strategies	8	47	0.71	48	0.65

Source: Hybrid Teacher Questionnaire and Control Teacher Questionnaire 2007/08 and 2008/09.

Confirmatory impact analysis

The statistical analyses for the two confirmatory research questions took the same general approach using a two-level hierarchical linear model, though the dependent variable and functional form of the estimating equations differed. For the first research question, the dependent variable was the student's test score on the pre-algebra/algebra section of the PLAN; for the second, it was a dichotomous indicator of whether grade 10 math course enrollment in the postintervention year was above algebra I. Using a hierarchical linear model accommodated the clustering of observations that were assumed to be present in the data (intraclass correlations), to ensure that standard errors were measured correctly. It allowed accounting for covariates that may be correlated with the treatment condition and with outcomes. Inclusion of these covariates in the model permitted obtaining unbiased estimates of the treatment effect and improved the precision of the estimates. Finally, the robustness of the study findings was tested by conducting sensitivity analyses. All models were specified before examining the outcome data.

Model for the PLAN assessment

The dependent variable for the first primary confirmatory analysis, PLAN, was a continuous variable defined as the student's scale score on the pre-algebra/algebra strand of the American College Testing PLAN: College Readiness Test for 10th Graders.

A two-level model with students nested within schools was estimated as follows:

Level 1 model (student level)

1. $Y_{ij} = \beta_{0j} + \beta_1 X_{1ij} + \ldots + \beta_8 X_{8ij} + r_{ij}$

where Y_{ij} is the value on the outcome (Y) for student i in school j, β_{0j} is the intercept, the X vector captures the student-level covariates and r_{ij} is the unmodeled residual for student i in school j.

Level 2 model (school level)

$$2.\quad \beta_{0j} = \gamma_0+\gamma_1 T_j+ \gamma_2 W_{1j} +\ldots+ \gamma_6 W_{5j} +\mu_{0j}$$

where β_{0j} is the school-level intercept, T_j represents the treatment condition (1) or control condition (0), W is the vector of school-level covariates, and μ_{0j} is the unmodeled residual for school j.

Evidence of impact was based on the statistical significance of the estimated parameter (γ1), using a two-tailed test with a 95 percent confidence interval and corrected for clustering at the school level. The coefficient (γ1) measured the impact of the offer of treatment on student outcomes.

The combined model is:

$$3.\quad Y_{ij} = \gamma_0+\gamma_1 T_j+ \gamma_2 W_{1j} +\ldots+ \gamma_6 W_{5j} + \beta_1 X_{1ij}+\ldots+ \beta_8 X_{8ij}+\mu_{0j} +r_{ij}$$

where the X vector captures the student-level covariates, T is the school-level treatment condition, and W is the vector of school-level covariates. Including these covariates improved the precision of the impact estimates and reduced the unexplained variation in outcomes between students in the treatment condition and those in the control condition. Although much of the difference among students had been eliminated through randomization, a cluster randomization design does not guarantee the elimination of all such differences with a small number of randomized units.

Model for grade 10 math course enrollment

The dependent variable for the second outcome in the primary confirmatory analysis—grade 10 math course enrollment—was an indicator of students' math attainment. The variable was a dichotomous indicator for whether the first math course a student enrolled in during the postintervention school year was above algebra I (0 = no, 1 = yes). The independent variables from equation 3 were used, but the dependent variable was replaced with this binary outcome, and the model was estimated using a logit link function. The following hierarchical generalized linear model was estimated:

$$4.\quad \eta_{ij} = \gamma_0+\gamma_1 T_j+ \gamma_2 W_{1j} +\ldots+ \gamma_6 W_{5j} + \beta_1 X_{1ij}+\ldots+ \beta_8 X_{8ij}+\mu_{0j}$$

where η_{ij} is the log of the odds of success (grade 10 enrollment in a math course above algebra I) for student i in school j, γ_0 is the school-level intercept, T_j represents the treatment condition (treatment [1] versus control [0]), and W is the vector for school-level variables. The X vector captures the baseline student-level characteristics listed previously. μ_{0j} is the unmodeled residual for school j.

Evidence of impact was based on the statistical significance of the estimated parameter (γ_1), using a two-tailed test with a 95 percent confidence interval and corrected for clustering at the school level. The sign of the coefficient (γ_1) indicated the direction of the impact of the offer of treatment on student outcomes. The results of the logistic regression were presented using the equation to calculate each student's probability of enrolling in a higher-level math course using individual student characteristics, and then generating group-level means for the treatment and control group using the treatment variable.

Student-level covariates

The models for both outcomes included these eight student-level covariates:

- Student-level deviation from the school-level average score on the math section of the grade 8 Kentucky Core Content Test (KCCT).
- Student-level deviation from the school-level average score on the math section of the grade 8 EXPLORE.
- Free and reduced-price lunch status.
- Gender.
- Underserved minority status.
- Individualized Education Plan status.
- Student's age in years.
- Course level indicating whether the student was enrolled in an honors-level algebra I course, rather than regular algebra I.

The student-level variables were included in this model to improve the precision of the impact estimates and to reduce the unexplained variation in outcomes between students in the treatment condition and those in the control condition. In particular, gender, free or reduced-price lunch status, and underserved minority status have all been shown in the empirical literature to correlate with math performance, after controlling for prior test score (Ginsburg-Block and Fantuzzo 1998; Mayer 1998; McCaffrey et al. 2001). Individualized Education Plan status represents the presence of a disability and is highly correlated with academic performance, while age acts as a proxy for grade level or course repetition, another indicator of prior academic performance. Students were assigned to either regular or honors algebra courses before the intervention. These variables were useful controls because different course levels may cover different amounts of content or cover content to different depths of understanding, and these factors may have affected student achievement.

School-level covariates

The models for both outcomes included these four school-level covariates:

- Rural school locale.
- School-level average score on the math section of the grade 8 KCCT.
- School-level average score on the math section of the grade 8 EXPLORE.
- Strata—a pair of dichotomous variables that represents how the school was selected in random assignment. In the exploratory analysis of whether the impact of the intervention differed by cohort, the strata variables were replaced with a single dichotomous variable for the cohort covariate.

Parameter to assess impact

The treatment variable (T) was dichotomous, indicating whether the student was enrolled in an algebra I course in the treatment group (T = 1) or the control group (T = 0). In the first primary confirmatory analysis, the coefficient γ_1 measured the impact of the intervention on student achievement, holding other variables in the model fixed. In the second primary confirmatory analysis, the coefficient γ_1 indicated the sign of the treatment effect on math attainment, holding other variables in the model fixed. For each impact estimate, the size of the impact was computed using Cohen's d. For the second research question, impacts were estimated using the regression equation to calculate each student's probability of enrolling in a higher-level math course using individual student characteristics, and then generating group-level means for the treatment and control groups using the treatment variable. The impact estimate was calculated by the first difference between the treatment and control groups.

Methodological issues

Methodological issues for the confirmatory impact analysis include missing data, sensitivity analyses, adjustments for multiple comparisons, and weighting.

Procedures for handling missing data

Missing data for dependent variables

Data were missing on the PLAN outcome for students who were not promoted, who left the Kentucky public school system, or who missed testing in grade 10. This means that results from this analysis apply only to students enrolled in algebra I at a participating school in grade 9, promoted to grade 10, and tested in grade 10. They do not generalize to algebra I students in participating schools who were not promoted to grade 10. No data were missing as a result of school attrition or untested classrooms, because the data came from Kentucky Department of Education administrative records, and because all grade 10 students were targeted for PLAN testing. There was no differential attrition between the treatment and control groups (response rates were 86 percent for the control group and 84 percent for the treatment group). Casewise deletion was used for records with missing values for the PLAN dependent variable. This method is appropriate when the outcome data are missing for students within schools in studies where the school is the unit of randomization (Puma et al. 2009).

No data were missing for the grade 10 math course enrollment outcome as a result of school attrition because the data were from Kentucky Department of Education administrative records. However, data on this outcome at the student level were missing because of incomplete administrative records and student attrition. In cohort I, 222 students (117 control and 105 treatment) or 6 percent were missing records for grade 10 math course enrollment. In cohort II, 277 students (129 control and 148 treatment) or 9 percent were missing records for grade 10 math course enrollment. These missing data were handled as follows:

- The grade 10 math enrollment data file lists all courses that students enrolled in during the postintervention year, but it does not include a separate indicator for students with no math course enrollments. As a result, students with a school enrollment record in the postintervention school year but missing a year 2 math enrollment record were assumed to be not taking any math course in year 2 and assigned a value of 0 for the dependent variable (*n* = 34 control students and 39 treatment students in cohort I; *n* = 24 control students and 40 treatment students in cohort II).
- Students missing year 2 math enrollment records were assumed to be lost to attrition if they had a withdrawal code at the beginning of the year indicating a transfer to an out-of-state, private, or home school; whereabouts unknown; or death. Casewise deletion was used for these records with missing values due to attrition (*n* = 14 control students and 9 treatment students in cohort I, *n* = 28 control students and 39 treatment students in cohort II).
- Students with no records in year 2 for school enrollment, math enrollment, or withdrawal status were assumed to be missing data because of incomplete administrative records. Casewise deletion was used for these records with missing values (*n* = 103 control students and 96 treatment students in cohort I, *n* = 102 control students and 110 treatment students in cohort II).

Missing data for independent variables

The dummy variable adjustment method was used for records with missing values for the independent variables. Using this approach, the value of the missing independent variables was set to a constant value of zero, and an additional dummy variable was added to the model to indicate whether the actual value was missing. This method is appropriate for dealing with missing baseline or pretest data in educational randomized controlled trials when data are missing for students within schools in studies that randomize at the school level (Puma et al. 2009).

Sensitivity analyses

The following sensitivity analyses were conducted to help establish the robustness of the main impact estimates:

1. *Estimating the models excluding the baseline covariates except for the student-level scale score pretests*.

This model provided an estimate of the difference in student outcomes between the treatment and control groups without controlling for any differences in student or school characteristics other than student pretest performance. This approach provided baseline estimates of impact of the intervention and determined whether the impact findings were sensitive to the model specification.

- Model (a) includes parameters associated only with student-level scale scores on the KCCT and EXPLORE pretests.
- Model (b) includes parameters associated only with student-level scale scores on the KCCT pretest.
- Model (c) includes parameters associated only with student-level scale scores on the EXPLORE pretest.

2. Estimating the full models with only one set of pretest variables at a time.

To determine whether the estimates were sensitive to the type and number of pretest variables included as baseline covariates in the full models, equations 3 and 4 were estimated again, excluding the variables for the EXPLORE school-level average pretest math score and EXPLORE student-level deviation and excluding the variables for the KCCT school-level average pretest math score and KCCT student-level deviation. Another reason for conducting this sensitivity analysis was that both the PLAN and EXPLORE are ACT-produced tests, and performance on EXPLORE should be a good predictor of performance on the PLAN. The Kentucky Department of Education expressed interest in learning whether, controlling for the EXPLORE score in math, the intervention affected PLAN outcomes.

3. Estimating the full models excluding students who were enrolled on September 1 in a part-year or part-credit course leading to completion of algebra I during the intervention period.

Kentucky schools offer algebra I using multiple schedules, some reducing exposure to the intervention for some study participants. The sample used in this sensitivity test excluded from the main confirmatory sample students not enrolled in a full-year algebra I course that began and ended within the intervention year. Specifically, it excluded students enrolled in only one semester of a course that used the intervention. Such students took the first semester of algebra I (part A) before the study began and took the second semester (part B) in the fall of the intervention year, or took the full-credit course

in a one-semester block-schedule format. In the part-year courses, students were exposed to the online courseware for a shorter period of time. If those part-year courses were taken in the fall semester, teachers also had less experience using the materials and participated in fewer school-year professional development sessions. In full-year part-credit courses, students would have used only the second half of the online course materials. Further, most part-year courses occur during the fall semester, and problems arose with the courseware during the fall semester for cohort I. In the fall of 2007, Kentucky Virtual Schools mistakenly loaded an old version of the student courseware, and it had errors. The correct version was uploaded early in the second semester of 2007/08. These conditions may have reduced the impact of the treatment for all participants involved in the intervention in 2007/08, but the problem was expected to be more severe for one-semester participants.

4. Estimating the full models excluding from cohort II the six duplicate schools that were in the control group in cohort I and then rerandomized in cohort II.

Using the two-cohort design, in which six schools in cohort II also participated in the cohort I control group, almost all student observations were independent.[36] One analytic concern about this design was that the randomized observations were no longer independent because of duplicate schools in both cohorts, which may have affected the estimated standard errors. The impact estimates were compared among the full sample in the confirmatory analysis and the sample from which the six duplicate schools were removed from cohort II in the sensitivity analysis, to determine if excluding the rerandomized schools (strata 2) affected the findings.

Adjustments for multiple comparisons

No adjustments were made for multiple comparisons because math achievement and math course-taking are not the same outcome domain.

Weighting

The data were not weighted. The number of algebra I students per school ranged from 20 to 373, and all those in grade 9 in each school were included in the study. Larger schools received more weight in the hierarchical linear model analysis because they had more students in the level 1 sample.

Exploratory impact analysis

Student subgroup effects

For exploratory research questions examining student subgroup effects, interaction terms were added to the hierarchical linear model used in the confirmatory analysis between the treatment variable and specific student covariate of interest. The first exploratory research question examined the difference in the impact of the intervention by gender and the second exploratory research question examined differences by cohort. The combined equation resulting from a two-level model with students nested within schools, and random effects associated with student characteristics, is shown below:

[36] There were nine students in the rerandomized schools who repeated the ninth grade, and thus were included in the analysis samples for both cohort I and cohort II.

5. $Y_{ij} = \gamma_0+\gamma_1 T_j+ \gamma_2 W_{1j} +\ldots+ \gamma_6 W_{5j} + \beta_1 X_{1ij}+\ldots+ \beta_8 X_{8ij}+\delta_{1j} T*X_{1ij} +\mu_{0j} + \mu_{1j} *X_{1ij}+r_{ij}$

where Y_{ij} is the outcome (Y) for student i in school j, γ_0 is the school-level intercept, T_j represents the treatment condition (0) versus control condition (1). W is a vector for level 2 variables representing rural status, strata (for the gender subgroup analysis) or cohort (for the cohort subgroup analysis), school-level average score on the math section of the grade 8 KCCT, and school-level average score on the math section of the grade 8 EXPLORE for algebra I students in school j. The X vector captures the baseline student-level covariates for student-level deviation from the school-level average score on the math section of the KCCT, student-level deviation from the school-level average score on the EXPLORE, free or reduced-price lunch status, gender, underserved minority status, Individualized Education Plan status, student's age in years, and course level. $T*X$ represents the interaction term between the treatment and the baseline student-level covariate of interest (gender or cohort I). μ_{0j} is the unmodeled residual for school j, and r_{ij} is the unmodeled residual for student i.

Equation 5 was modified to estimate a hierarchical generalized linear model for the grade 10 math course enrollment outcome to account for the change in the functional form of the dependent variable. The model was estimated as follows:

6. $\eta_{ij} = \gamma_0+\gamma_1 T_j+ \gamma_2 W_{1j} +\ldots+ \gamma_6 W_{5j} + \beta_1 X_{1ij}+\ldots+ \beta_8 X_{8ij}+\delta_{1j} T*X_{1ij} +\mu_{0j} + \mu_{1j} *X_{1ij}$,

where η_{ij} is the log of the odds of success for student i in school j, and all other covariates are the same as equation 5.

Parameter to assess impact

The treatment variable (T) was a dichotomous variable indicating whether the student was enrolled in an algebra I course in the treatment group (T = 1) or the control group (T = 0). For these research questions, the parameter for $T*X$ was examined, which represents the interaction term between the treatment and the baseline student-level covariates of interest, gender or cohort I.

The regression-adjusted means by gender and cohort were presented, based on the fully specified model with the interaction terms defined in the exploratory analysis. For the PLAN outcome, calculations were made by multiplying the coefficients in the regression model by the average values for the corresponding subgroup on each covariate. For the grade 10 math course enrollment outcome, impacts were estimated using the regression equation to calculate each student's probability of enrolling in a higher-level math course using individual student characteristics, and then generating group-level means for the treatment and control groups using the treatment variable. The impact estimate was the difference-in-difference, which is the difference between males in the treatment and control groups minus the difference between females in the treatment and control groups in the gender subgroup analysis. For the cohort subgroup analysis, the difference-in-difference is the difference between cohort I students in the treatment and

control groups minus the difference between cohort II students in the treatment and control groups.

School subgroup effects

For exploratory research questions estimating school subgroup effects, an interaction term was added to the hierarchical linear model used in the confirmatory analysis between the treatment variable and school covariates of interest. Student characteristics were treated as fixed effects, and treatment effects were allowed to vary with the school-level characteristic for whether the school is located in a rural or nonrural area. For the PLAN outcome variable, this two-level model was estimated as:

7. $Y_{ij} = \gamma_0+\gamma_1 T_j+ \gamma_2 W_{1j} +\ldots+ \gamma_6 W_{5j} + \beta_1 X_{1ij}+\ldots+ \beta_8 X_{8ij}+\delta_{1j} T{*}W_{1j}+\mu_{0j} + \mu_{1j} {*}W_{1j}+r_{ij}$

where $T{*}W$ represents the interaction terms between the treatment and the school-level characteristic of interest, rural status.

Equation 7 was modified to estimate a hierarchical generalized linear model for the grade 10 math course enrollment outcome to account for the change in the functional form of the model. The model was estimated as follows:

8. $\eta_{ij} = \gamma_0+\gamma_1 T_j+ \gamma_2 W_{1j} +\ldots+ \gamma_6 W_{5j} + \beta_1 X_{1ij}+\ldots+ \beta_8 X_{8ij}+\delta_{1j} T{*}W_{1j} +\mu_{0j} + \mu_{1j} {*}W_{1j}$

where η_{ij} is the log of the odds of success for student *i* in school *j*, and all other parameters are the same as equation 7.

Parameter to assess impact

The treatment variable (T) was dichotomous, indicating whether the student was enrolled in an algebra I course in the treatment group (T = 1) or the control group (T = 0). For these research questions, the parameter for $T{*}W$ was of interest. It represents the interaction term between the treatment and the baseline school-level characteristic of interest for rural status.

The regression-adjusted means by rural/nonrural locale were presented, based on the fully specified model with the interaction terms defined in the exploratory analysis. For the PLAN outcome, calculations were made by multiplying the coefficients in the regression model by the average values for the corresponding subgroup for rural status on each covariate. For the grade 10 math course enrollment outcome, impacts were estimated using the regression equation to calculate each student's probability of enrolling in a higher-level math course using individual student characteristics, and then generating group-level means for the treatment and control groups using the treatment variable. The impact estimate was the difference-in-difference, which is the difference between rural students in the treatment and control groups minus the difference between nonrural students in the treatment and control groups.

Methodological issues

The exploratory impact analyses used the same procedures as the confirmatory impact analyses for adjusting impact estimates for handling missing data, conducting sensitivity analyses, and weighting. No adjustments were made for multiple comparisons, since the analyses were exploratory.

Appendix E. Data Cleaning and File Construction

This appendix details the data-cleaning procedures for each source, the methods for dealing with problems in the data, and the extent to which problems in the data affected records in the treatment and control groups.

Student enrollment records

Student enrollment records provided by the Kentucky Department of Education were in long format, with multiple records for each course a student was enrolled in during the intervention school year. The records were sorted, based on the following order: student identification number, status on September 1 ("enrolled" records were sorted before "not enrolled" records), course start date, and course end date. Then a new data file was constructed, with a single record for each student containing information on the first course the student enrolled in on September 1. Under this method, a student enrolled in both a fall-semester and a spring-semester math course as of September 1 was assigned the record corresponding to the fall-semester course.

The course from the first enrollment record was used to determine whether a student was enrolled in an algebra I credit course eligible for inclusion in the sample. A separate set of enrollment variables were created, based on the last course of enrollment, to determine the extent to which students transferred to a different school and changed treatment or control status.

Methods for handling students concurrently enrolled in multiple math courses

The data were examined to determine if any records showed students enrolled concurrently in multiple math courses. (Concurrent enrollment was defined as enrollment in multiple math courses with start dates within 30 days of each other.) If a student was concurrently enrolled in an eligible and an ineligible course at the beginning of the school year, he or she was included in the sample with the enrollment record from the eligible course. Some students were concurrently enrolled in different levels of eligible courses (regular and honors); they were included in the sample as enrolled in an eligible course, but their level of course was categorized as missing.

In cohort I, 15 students (8 control, 7 treatment) were concurrently enrolled in more than one math course. None was enrolled in both an eligible and noneligible study course, so these concurrent records did not affect eligibility for inclusion in the sample.

In the cohort II enrollment file, 92 percent of the course start dates were before January 2009, suggesting that spring semester enrollment data were missing or incorrectly dated for most students. There was also no variable for the course end date in this file. For students with multiple enrollment records, this created some confusion about which record represented the initial course enrollment. In the control group, 70 students appeared to be concurrently enrolled in an eligible algebra I course and a noneligible algebra I lab. These students were included in the sample with the record from the eligible course. There were also three students in the control group and seven in the treatment group who appeared to be concurrently enrolled in different levels of eligible math courses (regular and honors). These students were coded as missing for the course-

level variable. The dummy variable adjustment method was used to indicate that the actual value was missing.

For less than 1 percent of the sample in cohort II, there were concurrent enrollment records on the same start date with different grade levels. The lowest grade level was selected, as this value should represent the status at the earliest point in the school year, if in fact the spring-semester enrollment records were incorrectly dated (*n* = 10 control students and 8 treatment students).

Missing enrollment data

Cohort I

The initial Kentucky Department of Education enrollment file for cohort I indicated students' enrollment status as of October 1. Because of concerns that students might have switched courses after the intervention started, a second data file with separate variables for enrollment status as of September 1 and October 1 was requested. The September 1 status date preceded the start of online course materials in study schools, making it a reference point preferable to October 1 since it was less likely to suffer from treatment-induced attrition. The two enrollment files were compared to determine if the number of students enrolled on October 1 was inconsistent with the September 1 status date. For 328 students (*n* = 196 control students and 132 treatment students), the October 1 enrollment records from the initial enrollment file were missing in the enrollment file with the September 1 status date. The enrollment records from the initial (October) file were used for these cases. Since the initial (October) enrollment file did not include a variable for enrollment status on September 1, it was assumed that the status was the same on September 1 and October 1. This seemed reasonable, as most students with data for both variables had the same value for both dates. In the randomized sample, 98 percent of the treatment group and 97 percent of the control group were enrolled on both September 1 and October 1. Approximately 1 percent of students (in both the treatment and control group) were enrolled on October 1 but not September 1; 1 percent of the control group and 2 percent of the treatment group were enrolled on September 1 but not October 1.

In addition, the Kentucky Department of Education enrollment records were compared with the end-of-course assessment records (supplemental data collected by the researchers at each study school, as described in appendix B) to determine if any classes were tested but had no enrollment records. For these instances, the schools provided enrollment records directly (*n* = 47 students).

Cohort II

The initial enrollment file received for cohort II appeared to be missing most enrollments. The Kentucky Department of Education was able to correct the problem and send a new file. The records from this new enrollment file were compared with the end-of-course assessment records to determine if any classes were tested but had no enrollment records. Students from five classes had end-of-course assessment records but were missing from the Kentucky Department of Education enrollment records. It was assumed that these students were also enrolled on September 1, and the information from the end-of-course assessment record was used as the enrollment record for the students in these classes (*n* = 42 control students and 36 treatment students). Approximately 90

percent of these students had pretest data from the prior year, suggesting that they were in grade 9 for the first time. Students missing a grade-level variable from the enrollment file were recoded as grade 9 if they had records in the pretest data (Kentucky Core Content Test [KCCT] or EXPLORE) in the year before the intervention (n = 40 control students and 32 treatment students).

Identifying noneligible enrollment records

In the combined cohorts, 9,069 students had an enrollment record in a study school during the intervention school year. Of these students, 76 percent (n = 6,908) were eligible for the sample. The remaining 2,161 noneligible records consisted of the following mutually exclusive categories:

- 218 students had initial enrollment records in a nonstudy school.
- 459 students had initial enrollment records in a study school but were not enrolled on September 1.
- 583 students were enrolled in a study school on September 1 but were not in an eligible algebra I credit course.
- 901 students were enrolled in a study school on September 1 in an eligible course but were not in grade 9.

Student demographic records

Demographic data with variables for gender, race, Individualized Education Plan status, free or reduced-priced lunch status, and age were collected from four sources (the Kentucky Department of Education, KCCT, school data, and EXPLORE), as described in chapter 3. Descriptive statistics (means, standard deviations, minimum values, maximum values) and tabulations were run for each variable, and no unusual or outlier values were found (table E1).

Table E1. Descriptive statistics for student-level demographics

Covariate	Number of student records with data	Mean across students	Standard deviation	Minimum	Maximum
Underserved minority (percent of students)	6,856	6.08	23.90	0.00	100.00
Male (percent of students)	6,857	51.70	49.97	0.00	100.00
Age (in years)	6,888	15.44	0.61	13.70	19.70
Recipient of free or reduced-price lunch (percent of students)	6,745	61.63	48.63	0.00	100.00
Enrolled in Individualized Education Plan (percent of students)	6,733	10.65	30.85	0.00	100.00

Source: Authors' calculations based on Kentucky Department of Education data for demographic characteristics (2007/08 and 2008/09), Kentucky Department of Education Commonwealth Accountability Testing System results for KCCT (2006/07 and 2007/08) and EXPLORE (2006/07 and 2007/08), and data provided directly by schools (2007/08 and 2008/09).

Inconsistencies among the demographic variables

The demographic file provided by the Kentucky Department of Education for cohort II had separate records for the fall and spring semesters. In almost all cases where the same student had multiple records, the value of the free or reduced-price lunch variable differed for each semester. Further, 98 percent of students in the sample were categorized as free or reduced-price lunch recipients in at least one of their demographic records. Given concerns about the accuracy of these data, this source was omitted when constructing the free or reduced-price lunch variable for cohort II. Instead, this variable was constructed using data from only the two pretests (KCCT and EXPLORE). In both these pretests, the Kentucky Department of Education provided the vendor with demographic data, including free or reduced-price lunch status, but the students could fill in a response to the demographic indicators on the test booklet, which would override the data sent by the Kentucky Department of Education.[37] For cohort II, pretest data were available for one and two years before the intervention, so this variable was available for both first-time and repeat grade 9 students who took the pretests. There appeared to be no problems with the free or reduced-price lunch variable in the Kentucky Department of Education demographic file for cohort I. Because pretest data were available for only one year before the intervention for this cohort, Kentucky Department of Education demographic data were used to construct the free or reduced-price lunch variable to reduce the missing data for repeat grade 9 students.

Given concerns about using separate data sources for the free or reduced-price lunch variable in cohorts I and II, researchers compared the match rate between the free or reduced-price lunch variables in the Kentucky Department of Education records and the pretest demographic records to determine the consistency of the coding for this variable. Cohort I had an 88 percent match between the free or reduced-price lunch variables in the Kentucky Department of Education and KCCT data files and an 89 percent match between the free or reduced-price lunch variables in the Kentucky Department of Education and EXPLORE data files among students with records in both files.

For cohort II, the Individualized Education Plan variable in the Kentucky Department of Education demographic file presented a problem, with 100 percent of records coded “no” for having an Individualized Education Plan. This source was omitted from the construction of the Individualized Education Plan variable and used data from KCCT as the sole data source for cohort II. There appeared to be no problems with the Individualized Education Plan variable in the demographic file for cohort I. The match rate between the Individualized Education Plan variables in the Kentucky Department of Education and KCCT files for students in cohort I with data from both sources was 97 percent.

Because of the issues with the free or reduced-price lunch and Individualized Education Plan variables from the Kentucky Department of Education in cohort II, researchers checked the other demographic variables very carefully and found no other problems. For students with demographic records in both the Kentucky Department of

[37] The Kentucky Department of Education transitioned to a new student information system after the 2007/08 pretests were administered. The problem with the free or reduced-price lunch variable occurred during the transfer of the data to the new system.

Education and KCCT files, the match rate between the two sources was 99 percent for race, 100 percent for gender, 100 percent for age (in years), and 99 percent for Individualized Education Plan .

Pretest records

The Kentucky Department of Education sent the data for the grade 8 KCCT and EXPLORE pretest for both cohorts. Some schools with missing enrollment records provided their KCCT records directly.

Missing pretest records

In cohort I, four records were missing from the Kentucky Department of Education KCCT file, and the data sent by the individual schools to the pretest records were added manually. Descriptive statistics (means, standard deviations, minimum values, maximum values) and tabulations were run for each variable, and no unusual or outlier values were found (table E2). All pretest values were within the range of possible test scores (800–880 for KCCT and 1–32 for EXPLORE).

Table E2. Descriptive statistics for pretest variables in the combined cohort

Pretest	Number	Mean	Standard deviation	Minimum	Maximum
KCCT math	6,266	837.17	19.30	800	880
EXPLORE math	6,073	13.85	3.17	2	25

Source: Authors' calculations based on Kentucky Department of Education Commonwealth Accountability Testing System results for KCCT (2006/07 and 2007/08) and EXPLORE (2006/07 and 2007/08).

Pretest records with duplicate student identification numbers

All students whose pretest records had missing or duplicate student identification numbers were dropped. In the pretest data for the entire state in cohort I (n = 50,743 for KCCT and n = 49,753 for EXPLORE), student identification numbers were missing for 425 records in the KCCT file and 287 records in the EXPLORE file. There were also records with duplicate student identification numbers, but different test scores, for 46 records in the KCCT file and 98 records in the EXPLORE file.

In the pretest data for the entire state in cohort II (n = 49,635 for KCCT and n = 48,410 for EXPLORE), student identification numbers were missing for 113 records in the KCCT file and 157 records in the EXPLORE file. There were also records with duplicate student identification numbers, but different test scores, for 40 records in the KCCT file and 109 records in the EXPLORE file.

PLAN records

The Kentucky Department of Education provided data for the PLAN outcome measure for the achievement analysis. Descriptive statistics (means, standard deviations, minimum values, maximum values) and tabulations were run for each variable, and no unusual or outlier values were found (table E3). All values were within the range of possible test scores (1–16 for the PLAN pre-algebra/algebra subscale).

Table E3. Descriptive statistics for pretest variables in the combined cohort

	Number	Mean	Standard deviation	Minimum	Maximum
PLAN pre-algebra/ algebra	5,863	7.03	2.81	1.00	16.00

Source: Authors' calculations based on Kentucky Department of Education Commonwealth Accountability Testing System results for PLAN (2008/09 and 2009/10).

PLAN records with duplicate student identification numbers

For the cohort I PLAN records for the entire state (*n* = 50,924), 592 records were missing student identification numbers and 134 records had duplicate identification numbers. Among the duplicates, 90 records had identical test dates but different scores, and 44 records had different test dates and different scores. Since students who transferred schools could have taken the PLAN more than once, the first PLAN record was examined for duplicate identification numbers with different test dates. All PLAN records statewide with missing or duplicate identification numbers with identical test dates were dropped. Multiple records with identical test dates but different scores were also dropped. Four identification numbers with more than one record (with different test dates and different scores) were included in the sample. In these cases, the student's record from the earliest test date was selected.

For the cohort II PLAN records for the entire state (*n* = 49,795), 608 were missing student identification numbers and 78 had duplicate identification numbers. Of the records with duplicates, 10 had different test dates and different scores, but none of these were included in the sample.

Grade 10 math course enrollment records and inconsistencies in the follow-up data files

The Kentucky Department of Education provided a math enrollment follow-up file tracking grade 9 students from cohort I year 1 into their year 2 math assignment (grade 10 for most students). It also provided separate files with data on school enrollment, withdrawal, and no-show records for the postintervention school year. In all the follow-up files, records were missing from four schools. Data were requested from administrators at each school.

Cross-tabulations among variables in the four follow-up data files (math enrollment, school enrollment, withdrawal, and no-show) were examined to identify any inconsistencies in students' enrollment status. The following changes were made:

- Seven students in cohort I and 30 students in cohort II (23 control and 14 treatment) had math follow-up records but no school enrollment records. These students were recoded as "enrolled" in the postintervention school year.
- Forty-four students in cohort I (25 control and 19 treatment) and 35 students in cohort II (14 control and 21 treatment) had withdrawal records but no school enrollment records. The enrollment status for these students was replaced with the reason for withdrawal.

- Thirteen students had records in the no-show file for the postintervention school year. All these students, however, also had a record in one of the other follow-up files. Students in the no-show file with school enrollment or math enrollment files were recoded as "enrolled." Students in the no-show file with withdrawal records were recoded with the reason for withdrawal as the enrollment status.

Inconsistencies in the coding of the follow-up files also affected the value of the outcome variable for grade 10 math course enrollment:

- Seventy-three students in cohort I (34 control and 39 treatment) and 64 students in cohort II (24 control and 40 treatment) had follow-up records in the school enrollment file but no math enrollment records. It was assumed that these students were not taking a math course in year 2 and each student was assigned the value 0 for the dependent variable of grade 10 math course enrollment.
- Seven students in cohort I (3 control and 4 treatment) and 34 students in cohort II (18 control and 16 treatment) with a withdrawal code indicating "dropout" at the beginning of the postintervention school year were either missing records in the year 2 math enrollment file or had reenrolled in a math course in the spring semester. These students were recoded as not enrolled in a grade 10 math course, to be consistent with the coding of other students categorized as dropouts at the beginning of the postintervention school year.
- Thirty-five students in cohort I had a withdrawal code indicating attrition at the beginning of the postintervention school year (transfer to an out-of-state, private, or home school; whereabouts unknown; or death). Of these students, 23 (14 control and 9 treatment) were missing data on grade 10 math enrollment and treated as missing data in the analysis of grade 10 math course-taking. In addition, 12 students (7 control and 5 treatment) had records in the year 2 math enrollment file, suggesting that they had registered for a math course before withdrawing from the school. For these 12 students, the record from the year 2 math enrollment file was used in the analysis of grade 10 math enrollment. For cohort II, 67 students (28 control and 39 treatment) were missing data on grade 10 math enrollment and treated as missing data in the analysis of grade 10 math course-taking. Also, eight students (three control and five treatment) had records in the year 2 math enrollment file, suggesting they had registered for a math course before withdrawing from the school. For these eight students, the record from the year 2 math enrollment file was used in the analysis of grade 10 math enrollment.

Several courses at individual study schools appeared to be miscoded with algebra I state course codes. Follow-up calls to the schools confirmed that these courses require algebra I as a prerequisite for enrollment. For students enrolled in these courses, the state codes in these records were recoded by the researchers to indicate grade 10 math enrollment in a course above algebra I.

Appendix F. Professional Development Timeline

Table F1 illustrates the hybrid algebra I professional development timeline. This appendix briefly describes each component below.

Table F1. Hybrid algebra I professional development timeline

May	June-August	September-May
One-day face-to-face • Orientation to hybrid study, meet learning community, and initial training in: o Blackboard o Kentucky Virtual Schools' algebra I o Spotlight on Algebra I	**1. Two-day face-to-face** **2. Five two-hour online sessions** **3. One-day face-to-face end-of-summer** • Horizon Wimba training • Complete Spotlight on Algebra I modules • Use of Blackboard course management system	**Monthly online session** • Facilitated learning community discussions regarding classroom implementation of hybrid algebra I

Source: Authors' compilation.

May

Professional development began in May with a one-day face-to-face session divided into three sections. The first section was a brief orientation, during which participants meet other teachers who form their community of learning. For the remainder of the morning, teachers learn to use the Blackboard platform (version 7.1) to navigate the Kentucky Virtual Schools' virtual classroom and are introduced to the Kentucky Virtual Schools' algebra I online course materials. During the second half of the day, an introductory unit of Spotlight on Algebra I gives teachers an overview of effective practices in a hybrid classroom.

For this study, treatment teachers also reviewed their responsibilities during the research period. The study paid for teachers' travel costs and paid schools for substitute teachers.

June–August

The summer professional development begins with a two-day face to-face session at the start of the summer break, during which teachers complete the next four units of the spotlight course. Training then shifts to an online format for five weeks and includes two-hour synchronous weekly sessions for all teachers, facilitated by the instructional specialists. Teachers are expected to read the units and answer written questions before each online meeting. Teachers who miss a meeting are expected to complete the assignment and read or listen to the recorded online session. Because Spotlight on Algebra I is an online/hybrid instructional program, it models the practices that teachers are to bring to their hybrid courses in algebra I. Teachers attended an end-of-summer face-to-face session to prepare them to register students in the online course and review guidelines for teaching hybrid algebra I with the National Repository of Online Courses courseware. The session also reviewed the use of Blackboard for course management, the professional development schedule, and the procedures for participating in the online

activities during the school year. Teachers received up to $600 compensation for participation in the summer professional development.

September–May

Monthly hour-long online sessions are scheduled after normal school hours to bring teachers together for additional guidance in using effective instructional practices in a hybrid environment. Multiple sessions are offered each month to meet the needs of teachers with busy, diverse schedules. Teachers must attend at least one session per month but may attend more often if they wish. These sessions, facilitated by the instructional specialists, schedule time for participants to share their experiences, successes, challenges, and solutions. The instructional specialists also visit classrooms in each treatment school once in the fall semester and up to three times total in the initial intervention year, to help teachers resolve implementation issues and improve their instructional strategies. These visits also provide the instructional specialists with formative information intended for use in adapting monthly online sessions to identified learning needs. During the study, teachers received up to $400 compensation for participation in the school-year professional development.

Appendix G. Detailed Teacher Survey Results

Table G1. Teacher survey results[38]

A. For combined cohorts

Item	Combined cohorts						
	Treatment (n=45)		Control (n=47)				
	Mean	Standard deviation	Mean	Standard deviation	Effect size	Mean difference	*p*
Items 1–16: Please indicate your level of agreement with the following statements by rating each one from: strongly disagree = 1 to strongly agree = 5. Note: items with "hybrid algebra 1 approach" were stated as "district algebra 1 approach" on control teacher surveys.							
1 Use of the hybrid algebra I approach is effective for helping students learn key algebraic concepts.	3.73	0.75	4.02	0.97	–0.34	–0.29*	0.03
2 My teaching is student-centered when I use the hybrid algebra I approach.	4.07	0.72	3.94	0.89	0.16	0.13	0.61
3 The hybrid algebra I approach emphasizes helpful learning activities.	3.91	0.73	3.38	0.99	0.61	0.53*	0.01
4 I think the hybrid algebra I approach effectively covers the knowledge and skills students need to successfully pass an algebra I end-of-course exam.	3.61	0.87	4.04	1.00	–0.46	–0.43*	0.01
5 Student interest and engagement are high when I use the hybrid algebra I approach.	3.20	0.97	3.04	0.88	0.17	0.16	0.40
6 The difficulty level of the KYVS algebra I student courseware is appropriate for most of my students.	3.29	1.01	3.66	0.96	–0.38	–0.37	0.06
7 I can implement the hybrid algebra I approach according to the recommended guidelines [as stated in the "note" included with the directions above].	3.87	1.04	3.96	0.67	–0.10	–0.09	0.81
8 I have received adequate training to effectively teach the hybrid algebra I approach.	4.16	0.98	4.17	0.96	–0.01	–0.01	0.95
9 I routinely use the hybrid algebra I approach.	4.07	0.96	4.40	0.71	–0.40	–0.33	0.10
10 I am able to align the KYVS algebra I approach with Kentucky's standards-based curriculum.	4.22	0.60	4.53	0.75	–0.46	–0.31*	< 0.01
11 The amount of academically focused class time is high when I use the hybrid algebra I approach.	3.64	0.86	4.02	0.77	–0.47	–0.38*	0.04
12 I have the essential algebra I knowledge and skills needed to conduct classes that implement the hybrid algebra I approach.	4.64	0.65	4.74	0.49	–0.18	–0.10	0.52
13 I have the essential technology knowledge and skills needed to conduct classes that implement the hybrid algebra I approach.	4.40	0.69	4.28	0.88	0.15	0.12	0.70
14 I can readily obtain answers to questions regarding implementation of the hybrid algebra I approach.	4.29	0.63	4.09	0.81	0.28	0.20	0.28

[38] Any inquiries regarding the teacher surveys or classroom observation instruments and methodology used for this project may be directed to: The Center for Research in Educational Policy, The University of Memphis, 325 Browning Hall, Memphis, TN 38152.

Item		Combined cohorts: Treatment (n=45) Mean	Treatment (n=45) Standard deviation	Control (n=47) Mean	Control (n=47) Standard deviation	Effect size	Mean difference	*p*
15	I often use computers to provide differentiated instruction based on individual learner needs.	3.51	0.91	2.60	1.22	0.85	0.91*	< 0.01
16	Use of the hybrid algebra I approach is effective for helping students learn key algebraic mechanics.	3.58	0.87	3.81	0.85	–0.27	–0.23	0.14
17	Ask "Why?" and "What if?" questions	4.09	0.51	4.19	0.61	–0.18	–0.10	0.35
18	Use number lines, graphs, or diagrams to explain algebra	4.27	0.58	4.49	0.62	–0.37	–0.22	0.06
19	Use a computer to explain algebra	3.71	0.89	3.00	1.14	0.70	0.71*	< 0.01
20	Work in groups	3.47	0.89	3.53	0.83	–0.07	–0.06	0.95
21	Write to explain algebra (descriptions, poetry, songs, reflections)	2.64	0.80	2.45	1.00	0.21	0.19	0.32
22	Talk to explain algebra	3.98	0.84	3.63	0.74	0.45	0.35*	0.01
23	Use things like algebra tiles or blocks	2.51	0.99	2.51	0.80	0.00	0.00	1.00
24	Use activities such as "guess and check," estimating, or drawing	3.47	0.99	3.28	0.88	0.21	0.19	0.32
25	Use graphing calculators	3.62	1.42	3.33	1.44	0.21	0.29	0.34
26	Use computers to learn algebra	3.76	0.98	1.93	0.74	2.13	1.83*	< 0.01
27	Use "exit slips"	2.82	1.17	2.35	1.17	0.41	0.47*	0.05
28	To what degree did the Hybrid/spotlight professional development change the way you teach algebra I? Scale: 1 = never, 2 = some, 3 = a lot	2.10	0.58	NA	NA	NA	NA	NA
Items 29–31: [treatment] Not counting hybrid and spotlight training, [treatment and control] to what degree has participation in the following math-related professional development/graduate work (completed during the past 12 months) changed the way you teach algebra I? Scale: 1 = never, 2 = some, 3 = a lot, NA = did not complete math professional development/courses.								
29	Workshops	2.02	0.60	2.26	0.66	–0.38	–0.24	0.09
30	Extended (nongraduate school) professional development programs	1.74	0.68	2.00	0.74	–0.37	–0.26	0.12
31	Graduate coursework	1.41	0.56	1.61	0.66	–0.33	–0.20	0.23

Note: Items 32–34 required open-ended responses.

Item		Combined cohorts: Treatment n=45 (yes) Number	Treatment n=45 (yes) Percent	Control n=47 (yes) Number	Control n=47 (yes) Percent	Difference Percent	χ^2	*p*
35	Would you like to teach a hybrid algebra I course again?	36	83.7	22	48.8	34.9*	12.97	<0.01

Dimension	Combined cohorts: Treatment (n=45) Mean	Treatment (n=45) Standard deviation	Control (n=47) Mean	Control (n=47) Standard deviation	Effect size	Mean difference	p
Impact on instruction (items 2,3,9,11,15)	3.84	0.58	3.67	0.63	0.28	0.17	0.17
Readiness to implement algebra I (items 7,8,10,12,13)	4.26	0.53	4.35	0.47	–0.18	–0.09	0.37
Impact on students (items 1, 4,5,6,16)	3.48	0.62	3.71	0.74	–0.34	–0.23*	0.04
Teacher use of algebra I strategies (items 17,18,19)	4.02	0.46	3.89	0.59	0.25	0.13	0.30
Student use of algebra I strategies (items 20–27)	3.28	0.61	2.73	0.38	1.06	0.55*	< 0.01

* Mean difference (from Wilcoxon-Mann-Whitney test) is statistically significant at the 95 percent confidence level using a two-tailed test.

NA is not applicable; item 28 did not apply to control teachers.

Note: Valid *n* for Hybrid Teacher Questionnaire = 45* treatment surveys and Control Teacher Questionnaire = 47 control surveys. Scale: 1 = never, 2 = rarely, 3 = occasionally, 4 = frequently, 5 = extensively. Effect sizes are calculated using Cohen's d with the pooled standard deviation. *Fewer than three teachers taught a fall semester Algebra class and a completely different spring semester Algebra class. These teachers completed two surveys – one for their fall class and a second for their spring class – and each of these surveys were included in the analysis.

Source: Authors' analysis based on the Hybrid Teacher Questionnaire and the Control Teacher Questionnaire (2007/08 and 2008/09).

B. For each cohort

Item		Cohort I: Treatment (n=21) Mean	Treatment (n=21) Standard deviation	Control (n=21) Mean	Control (n=21) Standard deviation	Effect size	Mean difference	p	Cohort II: Treatment (n=24) Mean	Treatment (n=24) Standard deviation	Control (n=26) Mean	Control (n=26) Standard deviation	Effect size	Mean difference	p
Items 1–16: Please indicate your level of agreement with the following statements by rating each one from: strongly disagree = 1 to strongly agree = 5. Note: items with "hybrid algebra 1 approach" were stated as "district algebra 1 approach" on control teacher surveys.															
1	Use of the hybrid algebra I approach is effective for helping students learn key algebraic concepts.	3.67	0.73	3.62	1.16	0.05	0.05	0.82	3.79	0.78	4.35	0.63	–0.81	–0.56*	0.01
2	My teaching is student-centered when I use the hybrid algebra I approach.	4.14	0.57	3.71	0.96	0.56	0.43	0.11	4.00	0.83	4.12	0.82	–0.15	–0.12	0.54
3	The hybrid algebra I approach emphasizes helpful learning	3.67	0.73	2.95	0.97	0.86	0.72*	0.01	4.13	0.68	3.73	0.87	0.52	0.40	0.12

	activities.														
4	I think the hybrid algebra I approach effectively covers the knowledge and skills students need to successfully pass an algebra I end-of-course exam.	3.70	0.86	3.67	1.15	0.03	0.03	0.89	3.54	0.88	4.35	0.75	–1.01	–0.81*	< 0.01
5	Student interest and engagement are high when I use the hybrid algebra I approach.	3.10	1.00	2.81	0.87	0.32	0.29	0.34	3.29	0.95	3.23	0.86	0.07	0.06	0.78
6	The difficulty level of the KYVS algebra I student courseware is appropriate for most of my students.	3.14	0.96	3.48	1.03	–0.35	–0.34	0.25	3.42	1.06	3.81	0.90	–0.41	–0.39	0.16
7	I can implement the hybrid algebra I approach according to the recommended guidelines [as stated in the “note” included with the directions above].	3.95	0.80	3.95	0.80	0.00	0.00	0.91	3.79	1.22	3.96	0.54	–0.19	–0.17	0.85
8	I have received adequate training to effectively teach the hybrid algebra I approach.	4.00	0.89	3.86	1.28	0.13	0.14	1.00	4.29	1.04	4.42	0.50	–0.16	–0.13	0.67
9	I routinely use the hybrid algebra I approach.	4.10	0.83	4.29	0.78	–0.24	–0.19	0.41	4.04	1.08	4.50	0.65	–0.53	–0.46	0.15
10	I am able to align the KYVS algebra I approach with Kentucky's standards-based curriculum.	4.29	0.56	4.29	0.96	0.00	0.00	0.53	4.17	0.64	4.73	0.45	–1.04	–0.56*	< 0.01
11	The amount of academically focused class time is high when I use the hybrid algebra I approach.	3.67	0.79	3.57	0.81	0.13	0.10	0.66	3.63	0.92	4.40	0.50	–1.07	–0.77*	<0 .01
12	I have the essential algebra I knowledge and skills needed to conduct classes that implement the hybrid algebra I approach.	4.71	0.46	4.76	0.44	–0.11	–0.05	0.74	4.58	0.78	4.73	0.53	–0.23	–0.15	0.59
13	I have the essential technology knowledge and skills needed to conduct classes that implement the hybrid algebra I approach.	4.24	0.77	4.05	0.97	0.22	0.19	0.61	4.54	0.59	4.46	0.76	0.12	0.08	0.89
14	I can readily obtain answers to questions regarding implementation of the hybrid	4.10	0.54	3.90	1.09	0.24	0.20	0.85	4.46	0.66	4.24	0.44	0.40	0.22	0.10

	algebra I approach.														
15	I often use computers to provide differentiated instruction based on individual learner needs.	3.42	1.02	2.52	1.33	0.77	0.90*	0.03	3.58	0.83	2.65	1.16	0.93	0.93*	< 0.01
16	Use of the hybrid algebra I approach is effective for helping students learn key algebraic mechanics.	3.62	0.80	3.48	0.98	0.16	0.14	0.63	3.54	0.93	4.08	0.63	–0.70	–0.54*	0.02
Items 17–19: [treatment] While implementing the hybrid algebra I approach this year…[control] While teaching algebra I this year how often did YOU do the following during DIRECT instruction***															
17	Ask “Why?” and “What if?” questions	4.19	0.60	4.10	0.54	0.10	0.09	0.56	4.00	0.42	4.27	0.67	–0.49	–0.27	0.08
18	Use number lines, graphs, or diagrams to explain algebra	4.33	0.48	4.38	0.67	–0.09	–0.05	0.63	4.21	0.66	4.58	0.58	–0.61	–0.37*	0.04
19	Use a computer to explain algebra	3.86	0.73	2.57	1.21	1.32	1.29*	< 0.01	3.58	1.02	3.35	0.98	0.23	0.23	0.42
Items 20–27: [treatment] While implementing the hybrid algebra I approach this year…[control] While teaching algebra I this past year how often did YOUR STUDENTS***															
20	Work in groups	3.57	0.68	3.10	0.70	0.70	0.47*	0.03	3.38	1.06	3.88	0.77	–0.55	–0.50	0.12
21	Write to explain algebra (descriptions, poetry, songs, reflections)	2.62	0.80	2.24	1.00	0.43	0.38	0.21	2.67	0.82	2.62	0.98	0.06	0.05	0.82
22	Talk to explain algebra	3.95	0.59	3.62	0.74	0.51	0.33	0.07	4.00	1.02	3.64	0.76	0.41	0.36	0.07
23	Use things like algebra tiles or blocks	2.76	0.94	2.24	0.83	0.60	0.52	0.07	2.29	1.00	2.73	0.72	–0.52	–0.44	0.09
24	Use activities such as “guess and check,” estimating, or drawing	3.57	0.98	3.10	1.00	0.49	0.47	0.16	3.38	1.01	3.42	0.76	–0.05	–0.04	0.98
25	Use graphing calculators	3.57	1.25	2.95	1.47	0.47	0.62	0.15	3.67	1.58	3.74	1.33	–0.05	–0.07	0.97
26	Use computers to learn algebra	3.86	0.73	1.90	0.77	2.68	1.96*	< 0.01	3.67	1.17	1.95	0.72	1.79	1.72*	<0 .01
27	Use “exit slips”	2.48	0.98	2.38	1.16	0.10	0.10	0.66	3.13	1.26	2.32	1.21	0.67	0.81*	0.02
28	To what degree did the hybrid/spotlight professional development change the way you teach algebra I? Scale: 1 = never, 2 = some, 3 = a lot	2.14	0.57	NA	NA	NA	NA	NA	2.05	0.59	NA	NA	NA	NA	NA

Items 29–31: [treatment] Not counting hybrid and spotlight training, [treatment and control] to what degree has participation in the following math-related professional development/graduate work (completed during the past 12 months) changed the way you teach algebra I? Scale: 1 = never, 2 = some, 3 = a lot, NA = did not complete math professional development/courses.														
29 Workshops	2.16	0.60	2.15	0.67	0.02	0.01	1.00	1.91	0.60	2.35	0.65	–0.72	–0.44*	0.02
30 Extended (nongraduate school) professional development programs	1.78	0.73	1.82	0.81	–0.05	–0.04	0.92	1.71	0.64	2.14	0.65	–0.69	–0.43*	0.04
31 Graduate coursework	1.36	0.50	1.57	0.65	–0.38	–0.21	0.40	1.45	0.60	1.63	0.68	–0.29	–0.18	0.40

	Cohort I							Cohort II						
	Treatment n=21 (yes)		Control n=21 (yes)		Difference	χ^2	*p*	Treatment n=24 (yes)		Control n=26 (yes)		Difference	χ^2	*p*
Item	**Number**	**Percent**	**Number**	**Percent**	**Percent**			**Number**	**Percent**	**Number**	**Percent**	**Percent**		
35 Would you like to teach a hybrid algebra I course again?	18	85.7	10	50.0	35.7*	6.93	0.01	18	81.8	12	48.0	33.8*	5.91	0.02

	Cohort I							Cohort II						
	Treatment (n=21)		Control (n=21)					Treatment (n=24)		Control (n=26)				
Dimension	**Mean**	**Standard deviation**	**Mean**	**Standard deviation**	**Effect size**	**Mean difference**	***p***	**Mean**	**Standard deviation**	**Mean**	**Standard deviation**	**Effect size**	**Mean difference**	***p***
Impact on instruction (items 2,3,9,11,15)	3.80	0.47	3.41	0.66	0.69	0.39*	0.04	3.88	0.67	3.88	0.52	0.00	0.00	0.83
Readiness to implement algebra I (items 7,8,10,12,13)	4.24	0.50	4.18	0.54	0.12	0.06	0.81	4.28	0.56	4.46	0.40	–0.38	–0.18	0.19
Impact on students (items 1, 4,5,6,16)	3.43	0.58	3.40	0.87	0.04	0.03	0.76	3.52	0.67	3.96	0.52	–0.75	–0.44*	0.01
Teacher use of algebra I strategies (items 17,18,19)	4.13	0.37	3.68	0.54	1.00	0.45*	< 0.01	3.93	0.52	4.06	0.57	–0.24	–0.13	0.40
Student use of algebra I strategies (items 20–27)	3.30	0.39	2.69	0.39	1.62	0.61*	< 0.01	3.27	0.76	3.07	0.60	0.30	0.20	0.14

* Mean difference (from Wilcoxon-Mann-Whitney test) is statistically significant at the 95 percent confidence level using a two-tailed test.

NA is not applicable; item 28 did not apply to control teachers.

Note: Valid *n* for Hybrid Teacher Questionnaire = 45* treatment surveys and Control Teacher Questionnaire = 47 control surveys. Scale: 1 = never, 2 = rarely, 3 = occasionally, 4 = frequently, 5 = extensively. Effect sizes are calculated using Cohen's d with the pooled standard deviation. *Fewer than three teachers taught a fall semester Algebra class and a completely different spring semester Algebra class. These teachers completed two surveys – one for their fall class and a second for their spring class – and each of these surveys were included in the analysis.

Source: Authors' analysis based on the Hybrid Teacher Questionnaire and the Control Teacher Questionnaire (2007/08 and 2008/09).

Table G2. School Observation Measure results

A. For combined cohorts

Item		Combined cohorts						
		Treatment (*n* = 80)		Control (*n* = 85)				
		Mean	Standard deviation	Mean	Standard deviation	Effect size	Mean difference	*p*
1	Direct instruction (lecture)	2.53	1.39	3.07	0.84	–0.48	–0.54*	0.03
2	Cooperative/collaborative learning	0.63	1.10	0.78	1.26	–0.13	–0.15	0.54
3	Higher-level instructional feedback (written or verbal) to enhance student learning	1.28	1.29	0.80	1.21	0.39	0.48*	0.01
4	Use of higher-level questioning strategies	1.41	1.25	1.18	1.26	0.18	0.23	0.19
5	Teacher acting as a coach/facilitator	2.21	1.52	2.34	1.30	–0.09	–0.13	0.78
6	Independent seatwork (self-paced worksheets, individual assignments)	0.98	1.27	2.22	1.14	–1.04	–1.24*	< 0.01
7	Student discussion	0.30	0.79	0.12	0.47	0.28	0.18	0.06
8	Computer for instructional delivery (computer-assisted instruction, drill and practice)	2.26	1.76	1.27	1.59	0.59	0.99*	< 0.01
9	Technology as a learning tool or resource (Internet research, spreadsheet or database creation, multimedia, CD-ROM, Laserdisc)	0.75	1.39	0.45	1.07	0.24	0.30	0.19
10	High academically focused class time	3.39	0.74	3.42	0.93	–0.04	–0.03	0.27
11	High level of student attention/interest/engagement	3.09	0.84	2.78	1.02	0.33	0.31*	0.05

* Mean difference (from Wilcoxon-Mann-Whitney test) is statistically significant at the 95 percent confidence level using a two-tailed test.

Note: Scale: 0 = not observed, 1 = rarely, 2 = occasionally, 3 = frequently, 4 = extensively. Effect sizes are calculated using Cohen's d with the pooled standard deviation.

Source: Authors' analysis based on the School Observation Measure (2007/08 and 2008/09).

B. For each cohort

Item	Cohort I							Cohort II						
	Treatment (*n* = 38)		Control (*n* = 36)					Treatment (*n* = 42)		Control (*n* = 49)				
	Mean	Standard deviation	Mean	Standard deviation	Effect size	Mean difference	*p*	Mean	Standard deviation	Mean	Standard deviation	Effect size	Mean difference	*p*
1 Direct instruction (lecture)	2.87	1.28	2.97	1.00	–0.09	–0.10	0.95	2.21	1.42	3.14	0.71	–0.86	–0.93*	< 0.01
2 Cooperative/collaborative learning	0.24	0.79	0.89	1.30	–0.62	–0.65*	0.01	0.98	1.22	0.69	1.23	0.24	0.29	0.13
3 Higher-level instructional feedback (written or verbal) to enhance student learning	2.08	1.22	0.31	0.79	1.74	1.77*	< 0.01	0.55	0.86	1.16	1.34	–0.54	–0.61*	0.02
4 Use of higher-level questioning strategies	2.00	1.27	0.72	0.94	1.16	1.28*	< 0.01	0.88	0.97	1.51	1.36	–0.53	–0.63*	0.03
5 Teacher acting as a coach/facilitator	1.79	1.63	2.14	1.33	–0.24	–0.35	0.37	2.60	1.33	2.49	1.26	0.09	0.11	0.59
6 Independent seatwork (self-paced worksheets, individual assignments)	1.79	1.23	2.00	1.20	–0.18	–0.21	0.36	0.24	0.76	2.39	1.08	–2.30	–2.15*	< 0.01
7 Student discussion	0.13	0.53	0.14	0.42	–0.02	–0.01	0.66	0.45	0.94	0.10	0.51	0.48	0.35*	0.01
8 Computer for instructional delivery (computer-assisted instruction, drill and practice)	2.18	1.84	0.94	1.51	0.74	1.24*	< 0.01	2.33	1.71	1.51	1.62	0.50	0.82*	0.02
9 Technology as a learning tool or resource (Internet research, spreadsheet or database creation, multimedia, CD-ROM, Laserdisc)	0.68	1.40	0.78	1.38	–0.07	–0.10	0.61	0.81	1.40	0.20	0.71	0.57	0.61*	0.02
10 High academically focused class time	3.32	0.77	3.06	1.04	0.29	0.26	0.35	3.45	0.71	3.69	0.74	–0.33	–0.24*	0.02
11 High level of student attention/interest/engagement	2.84	0.82	2.42	1.16	0.43	0.42	0.09	3.31	0.81	3.04	0.82	0.33	0.27	0.08

* Mean difference (from Wilcoxon-Mann-Whitney test) is statistically significant at the 95 percent confidence level using a two-tailed test.
Note: Scale: 0 = not observed, 1 = rarely, 2 = occasionally, 3 = frequently, 4 = extensively. Effect sizes are calculated using Cohen’s d with the pooled standard deviation.
Source: Authors’ analysis based on the School Observation Measure (2007/08 and 2008/09).

Table G3. Algebra I Quality Assessment results: observed activities

A. For combined cohorts

	Item	Combined cohorts: Treatment (n = 80) Number	Treatment (n = 80) Percent	Control (n = 85) Number	Control (n = 85) Percent	Difference Percent	χ^2	p
	Teacher activities							
1	Ask “Why?” and “What if?” questions	56	70.0	50	58.8	11.2	2.24	0.13
2	Use number lines, graphs, or diagrams to explain algebra	51	63.8	54	63.5	0.30	< 0.01	0.98
3	Use a computer to explain algebra	28	35.0	32	37.7	–2.7	0.13	0.72
	Student activities							
4	Work in groups	19	24.1	33	38.8	–14.7*	4.13	0.04
5	Write to explain algebra (descriptions, poetry, songs, reflections)	4	5.1	8	9.4	–4.3	1.14	0.29
6	Talk to explain algebra	32	40.5	32	37.7	2.8	0.14	0.71
7	Use things like algebra tiles or blocks	4	5.1	6	7.1	–2.0	0.29	0.59
8	Use activities such as “guess and check,” estimating, or drawing	28	35.4	13	15.3	20.1*	8.87	< 0.01
9	Use graphing calculators	29	36.7	38	56.7	–20.0	1.08	0.30
10	Use computers to learn algebra	40	50.6	9	10.6	40.0*	31.34	< 0.01
11	Use “exit slips”	5	6.3	7	8.2	50.6	0.22	0.64

* Difference (from chi-square test) is statistically significant at the 95 percent confidence level using a two-tailed test.

Note: Scale: 0 = no, 1 = yes.

Source: Authors’ analysis based on the Algebra I Quality Assessment (2007/08 and 2008/09).

B. For each cohort

Item	Cohort I							Cohort II						
	Treatment (n = 38)		Control (n = 36)		Difference	χ^2	p	Treatment (n = 42)		Control (n = 49)		Difference	χ^2	p
	Number	Percent	Number	Percent	Percent			Number	Percent	Number	Percent	Percent		
Teacher activities														
1 Ask “Why?” and “What if?” questions	34	89.5	14	38.9	50.6*	20.76	< .01	22	52.4	36	73.5	–21.1*	4.35	0.04
2 Use number lines, graphs, or diagrams to explain algebra	27	71.1	18	50.0	21.1	3.44	.06	24	57.1	36	73.5	–16.4	2.68	0.10
3 Use a computer to explain algebra	14	36.8	6	16.7	20.1	3.82	.05	14	33.3	26	53.1	–19.8	3.57	0.06
Student activities														
4 Work in groups	6	16.2	13	36.1	–19.9	3.75	.05	13	31.0	20	40.8	–9.8	0.95	0.33
5 Write to explain algebra (descriptions, poetry, songs, reflections)	3	8.1	‡	‡	‡	‡	‡	‡	‡	7	14.3	‡	‡	‡
6 Talk to explain algebra	20	54.1	15	41.7	12.4	1.12	.29	12	28.6	17	34.7	–6.1	0.39	0.53
7 Use things like algebra tiles or blocks	3	8.1	4	11.1	–3.0	0.19	.66	‡	‡	3	5.8	‡	‡	‡
8 Use activities such as “guess and check,” estimating, or drawing	13	35.1	10	27.8	7.3	0.46	.50	15	35.7	3	6.1	29.6*	12.48	< 0.01
9 Use graphing calculators	10	27.0	17	47.2	–20.2	3.19	.07	19	45.2	21	42.9	2.3	0.05	0.82
10 Use computers to learn algebra	15	40.5	7	19.4	21.1*	3.86	.05	25	59.5	‡	‡	‡	‡	‡
11 Use “exit slips”	3	8.1	‡	‡	‡	‡	‡	‡	‡	5	10.2	‡	‡	‡

* Difference (from chi-square test) is statistically significant at the 95 percent confidence level using a two-tailed test.
‡ Reporting standards not met (n < 3).
Note: Scale: 0 = no, 1 = yes.
Source: Authors’ analysis based on the Algebra I Quality Assessment (2007/08 and 2008/09).

Table G4. Algebra I Quality Assessment results: quality of observed activities

A. For combined cohorts

Item	Combined cohorts: Treatment (*n* = 80) Mean	Treatment Standard deviation	Control (*n* = 85) Mean	Control Standard deviation	Effect size	Mean difference	*p*
Teacher activities							
1 Ask "Why?" and "What if?" questions	2.22	0.69	2.19	0.82	0.04	0.03	0.97
2 Use number lines, graphs, or diagrams to explain algebra	2.24	0.62	2.30	0.69	–0.09	–0.06	0.57
3 Use a computer to explain algebra	1.93	0.84	2.53	0.72	–0.78	–0.60*	< 0.01
Student activities							
4 Work in groups	2.37	0.76	2.31	0.64	0.09	0.06	0.65
5 Write to explain algebra (descriptions, poetry, songs, reflections)	2.50	1.00	1.75	0.46	1.23	0.75	0.11
6 Talk to explain algebra	1.94	0.72	2.16	0.72	–0.31	–0.22	0.22
7 Use things like algebra tiles or blocks	1.60	0.89	2.17	0.75	–0.77	–0.57	0.28
8 Use activities such as "guess and check," estimating, or drawing	2.29	0.71	1.31	0.48	1.55	0.98*	< 0.01
9 Use graphing calculators	2.64	0.49	2.16	0.75	0.75	0.48*	< 0.01
10 Use computers to learn algebra	2.83	0.44	2.33	0.87	0.95	0.50*	0.03
11 Use "exit slips"	2.20	1.10	2.71	0.49	–0.70	–0.51	0.49

* Mean difference (from Wilcoxon-Mann-Whitney test) is statistically significant at the 95 percent confidence level using a two-tailed test.
Note: Scale: 1 = low, 2 = moderate, 3 = high. Effect sizes are calculated using Cohen's d with the pooled standard deviation.
Source: Authors' analysis based on the Algebra I Quality Assessment (2007/08 and 2008/09).

B. For each cohort

Item	Cohort I							Cohort II						
	Treatment (n = 38)		Control (n = 36)					Treatment (n = 42)		Control (n = 49)				
	Mean	Standard deviation	Mean	Standard deviation	Effect size	Mean difference	p	Mean	Standard deviation	Mean	Standard deviation	Effect size	Mean difference	p
Teacher activities														
1 Ask “Why?” and “What if?” questions	2.03	0.72	1.67	0.79	0.50	0.36	0.14	2.52	0.51	2.36	0.76	0.24	0.16	0.61
2 Use number lines, graphs, or diagrams to explain algebra	2.11	0.64	1.72	0.57	0.65	0.39*	0.05	2.39	0.58	2.58	0.55	–0.34	–0.19	0.19
3 Use a computer to explain algebra	2.14	0.77	1.83	0.98	0.39	0.31	0.48	1.73	0.88	2.69	0.55	–1.43	–0.96*	< 0.01
Student activities														
4 Work in groups	1.67	0.52	2.00	0.71	–0.53	–0.33	0.35	2.69	0.63	2.53	0.51	0.29	0.16	0.26
5 Write to explain algebra (descriptions, poetry, songs, reflections)	2.33	1.15	‡	‡	‡	‡	‡	‡	‡	1.86	0.38	‡	‡	‡
6 Talk to explain algebra	1.90	0.79	1.80	0.77	0.13	0.10	0.72	2.00	0.60	2.47	0.51	–0.89	–0.47*	0.05
7 Use things like algebra tiles or blocks	1.75	0.96	1.75	0.50	0.00	0.00	1.00	‡	‡	‡	‡	‡	‡	‡
8 Use activities such as “guess and check,” estimating, or drawing	1.85	0.56	1.20	0.42	1.36	0.65*	0.01	2.67	0.62	1.67	0.58	1.72	1.00*	0.02
9 Use graphing calculators	2.44	0.53	1.65	0.61	1.42	0.79*	<0.01	2.74	0.45	2.57	0.60	0.33	0.17	0.40
10 Use computers to learn algebra	2.63	0.62	2.71	0.49	–0.14	–0.08	0.89	2.96	0.20	‡	‡	‡	‡	‡
11 Use “exit slips”	2.20	1.10	‡	‡	‡	‡	‡	‡	‡	3.00	0.00	‡	‡	‡

* Mean difference (from Wilcoxon-Mann-Whitney test) is statistically significant at the 95 percent confidence level using a two-tailed test.
‡ Reporting standards not met (n < 3).
Note: Scale: 1 = low, 2 = moderate, 3 = high. Effect sizes are calculated using Cohen’s d with the pooled standard deviation.
Source: Authors’ analysis based on the Algebra I Quality Assessment (2007/08 and 2008/09).

Table G5. Algebra I Quality Assessment results: mean numbers of observed teacher and student activities

	Teacher activities										Student activities									
	Treatment			Control			Mean difference	*t*	*p*	Effect size	Treatment			Control			Mean difference	*t*	*p*	Effect size
	n	**M**	**SD**	***n***	**M**	**SD**					***n***	**M**	**SD**	***n***	**M**	**SD**				
Cohort I	38	1.97	0.94	36	1.06	0.67	0.91*	–4.79	<0 .01	1.13	38	1.92	1.40	36	1.92	1.44	0.00	–0.01	0.99	0.00
Cohort II	42	1.43	1.02	49	2.00	0.89	–0.57*	2.86	0.01	–0.61	42	1.43	1.02	49	1.57	1.50	–0.14	–1.92	0.06	–0.11
Total	80	1.69	1.01	85	1.60	0.93	0.09	–0.58	0.56	0.09	80	2.01	1.25	85	1.72	1.48	0.29	–1.38	0.17	0.21

* Mean difference (from t- test) is statistically significant at the 95 percent confidence level using a two-tailed test.
M = mean; SD = standard deviation.
Note: Effect sizes are calculated using Cohen’s d with the pooled standard deviation.
Source: Authors’ analysis based on the Algebra I Quality Assessment (2007/08 and 2008/09).

Appendix H. Results of Sensitivity Analyses

Four sets of sensitivity analyses were conducted to further establish the robustness of the main impact findings. This appendix presents the results of these analyses.

Description of sensitivity analysis models

This section describes each model run for the sensitivity analyses.

1. Estimating the models excluding the baseline covariates except for the student-level scale score pretests.

- Model 1(a) includes parameters associated only with student-level scale scores on the Kentucky Core Content Test (KCCT) and EXPLORE pretests.
- Model 1(b) includes parameters associated only with student-level scale scores on the KCCT pretest.
- Model 1(c) includes parameters associated only with student-level scale scores on the EXPLORE pretest.

2. Estimating the full models with only one set of pretest variables at a time.

- Model 2(a) includes all covariates from the full model except for the EXPLORE pretest.
- Model 2(b) includes all covariates from the full model except for the KCCT pretest.

3. Estimating the full models excluding students enrolled on September 1 in a part-year or part-credit course leading to completion of algebra I credit during the intervention period.

4. Estimating the full models excluding from cohort II the six duplicate schools that were in the control group in cohort I and then rerandomized in cohort II.

Results of sensitivity analyses

Table H1 presents the results from each set of sensitivity analyses for the confirmatory impact analysis on student achievement on the grade 10 pre-algebra/algebra PLAN.

Table H1. Sensitivity analysis of confirmatory impact findings for student achievement results in grade 10 pre-algebra/algebra on the PLAN assessment

Type of sensitivity analysis	Treatment group mean (standard deviation)	Control group mean (standard deviation)	Estimated impact (standard error)	*p*	95 percent confidence interval	Estimated impact in effect-size units[a]
1a) PLAN outcome adjusted for pretest variables only	6.81 (2.68)	7.10 (2.91)	–0.28 (0.17)	0.10	–0.62, 0.06	–0.10
1b) PLAN outcome adjusted for KCCT pretest variables only	6.82 (2.68)	7.11 (2.91)	–0.29 (0.19)	0.13	–0.67, 0.08	–0.10
1c) PLAN outcome adjusted for EXPLORE pretest variables only	6.79 (2.68)	7.10 (2.91)	–0.31 (0.17)	0.07	–0.66, 0.03	–0.11
2a) PLAN outcome adjusted for student and school characteristics *except* EXPLORE pretest	6.83 (2.68)	7.10 (2.91)	–0.27 (0.18)	0.15	–0.63, 0.10	–0.09
2b) PLAN outcome adjusted for student and school characteristics *except* KCCT pretest	6.81 (2.68)	7.06 (2.91)	–0.25 (0.20)	0.21	–0.64, 0.14	–0.09
3) PLAN outcome adjusted for student and school characteristics, excluding students enrolled in a part-year or part-credit course	6.76 (2.68)	6.83 (2.91)	–0.07 (0.17)	0.67	–0.40, 0.26	–0.03
4) PLAN outcome adjusted for student and school characteristics, excluding the six duplicate schools in cohort I that were rerandomized in cohort II	6.89 (2.68)	7.10 (2.91)	–0.21 (0.20)	0.29	–0.60, 0.18	–0.08

a. Effect sizes are calculated using Cohen's d with the pooled standard deviation.
* Coefficient (of estimated intent-to-treat impact) is statistically significant at the 95 percent confidence level using a two-tailed test.
Note: Valid *n* for PLAN = 5,863.
Source: Authors' analysis based on Kentucky Department of Education data for enrollment (2007/08 and 2008/09) and demographic characteristics (2007/08 and 2008/09), Kentucky Department of Education Commonwealth Accountability Testing results for KCCT (2006/07 and 2007/08), EXPLORE (2006/07 and 2007/08), and PLAN (2007/08 and 2008/09), and data provided directly by schools (2007/08 and 2008/09).

Table H2 presents the sensitivity analysis of the confirmatory impact findings on grade 10 math course enrollment.

Table H2. Sensitivity analysis of confirmatory impact findings for grade 10 math course enrollment

Type of sensitivity analysis	Treatment group probability	Control group probability	Difference[a]	Odds ratio (standard error)	Log odds ratio	*p*
1a) Grade 10 math enrollment outcome adjusted for pretest variables only	0.87	0.87	0.00	0.97 (0.32)	–0.03	0.92
1b) Grade 10 math enrollment outcome adjusted for KCCT pretest variables only	0.87	0.87	0.00	0.97 (0.32)	–0.03	0.94
1c) Grade 10 math enrollment outcome adjusted for EXPLORE pretest variables only	0.87	0.87	0.00	0.94 (0.31)	–0.06	0.86
2a) Grade 10 math enrollment outcome adjusted for student and school characteristics *except* EXPLORE pretest	0.86	0.87	–0.01	1.31 (0.43)	0.27	0.41
2b) Grade 10 math enrollment outcome adjusted for student and school characteristics *except* KCCT pretest	0.86	0.86	0.00	1.34 (0.43)	0.29	0.36
3) Grade 10 math enrollment outcome adjusted for student and school characteristics, excluding students enrolled in a part-year or part-credit course	0.86	0.86	0.00	1.26 (0.45)	0.23	0.51
4) Grade 10 math enrollment outcome adjusted for student and school characteristics, excluding the six duplicate schools in cohort I that were rerandomized in cohort II	0.85	0.86	–0.01	1.20 (0.42)	0.18	0.61

a. Impact estimates are estimated using the regression equation to calculate each student's probability of enrolling in a higher-level math course using individual student characteristics, and then generating group-level means for the treatment and control groups using the treatment variable. The impact estimate is calculated as the first difference between treatment and control groups.
* Difference (of estimated intent-to-treat impact) is statistically significant at the 95 percent confidence level using a two-tailed test.
Note: Valid *n* for grade 10 math course enrollment = 6,409.
Source: Authors' analysis based on Kentucky Department of Education data for enrollment (2007/08, 2008/09, and 2009/10) and demographic characteristics (2007/08 and 2008/09), Kentucky Department of Education Commonwealth Accountability Testing System results for KCCT (2006/07 and 2007/08) and EXPLORE (2006/07 and 2007/08), and data provided directly by schools (2007/08 and 2008/09).

Tables H3, H4, and H5 present the results from the sensitivity analyses for the exploratory impact analysis on student achievement on the grade 10 pre-algebra/algebra PLAN.

Table H3. Sensitivity analysis of exploratory impact findings for student achievement results in grade 10 pre-algebra/algebra on the PLAN assessment, by gender

Type of sensitivity analysis	Student subgroup	Treatment group mean (standard deviation)	Control group mean (standard deviation)	Difference (Standard error)	Estimated impact in effect-size units[a]	*p*
1a) PLAN outcome adjusted for pretest variables only	Male	6.77 (2.87)	7.01 (3.06)	–0.24	–0.08	
	Female	6.55 (2.47)	6.89 (2.76)	–0.34	–0.13	
	Difference	0.22 (2.68)	0.12 (2.91)	0.10 (0.11)	0.04	0.39
1b) PLAN outcome adjusted for KCCT pretest variables only	Male	6.76 (2.87)	6.99 (3.06)	–0.24	–0.08	
	Female	6.59 (2.47)	6.94 (2.76)	–0.35	–0.13	
	Difference	0.17 (2.68)	0.05 (2.91)	0.11 (0.12)	0.04	0.34
1c) PLAN outcome adjusted for EXPLORE pretest variables only	Male	6.83 (2.87)	7.09 (3.06)	–0.26	–0.09	
	Female	6.52 (2.47)	6.89 (2.76)	–0.38	–0.15	
	Difference	0.31 (2.68)	0.20 (2.91)	0.11 (0.12)	0.04	0.35
2a) PLAN outcome adjusted for student and school characteristics *except* EXPLORE pretest	Male	6.82 (2.87)	7.01 (3.06)	–0.19	–0.06	
	Female	6.54 (2.47)	6.87 (2.76)	–0.34	–0.13	
	Difference	0.28 (2.68)	0.14 (2.91)	0.14 (0.11)	0.05	0.21
2b) PLAN outcome adjusted for student and school characteristics *except* KCCT pretest	Male	6.88 (2.87)	7.06 (3.06)	–0.18	–0.06	
	Female	6.48 (2.47)	6.80 (2.76)	–0.32	–0.12	
	Difference	0.40 (2.68)	0.26 (2.91)	0.14 (0.12)	0.05	0.24
3) PLAN outcome adjusted for student and school characteristics, excluding students enrolled in a part-year or part-credit course	Male	6.81 (2.87)	6.97 (3.06)	–0.16	–0.05	
	Female	6.52 (2.47)	6.82 (2.76)	–0.29	–0.11	
	Difference	0.29 (2.68)	0.15 (2.91)	0.14 (0.12)	0.05	0.24
4) PLAN outcome adjusted	Male	6.85	6.98	–0.13	–0.04	

Type of sensitivity analysis	Student subgroup	Treatment group mean (standard deviation)	Control group mean (standard deviation)	Difference (Standard error)	Estimated impact in effect-size units[a]	*p*
for student and school characteristics, excluding the six duplicate schools in cohort I that were rerandomized in cohort II		(2.87)	(3.06)			
	Female	6.52 (2.47)	6.81 (2.76)	–0.29	–0.11	
	Difference	0.33 (2.68)	0.18 (2.91)	0.16 (0.12)	0.06	0.19

a. Effect sizes are calculated using Cohen's d with the pooled standard deviation.
* Difference (of estimated intent-to-treat impact) is statistically significant at the 95 percent confidence level using a two-tailed test.
Note: Means are regression adjusted using the same set of baseline covariates as the confirmatory analysis. Valid *n* for PLAN = 5,863, including: (male = 2,951, female = 2,900, and missing gender =12). The *p*-value is for the coefficient on the interaction term between treatment and male.
Source: Authors' analysis based on Kentucky Department of Education data for enrollment (2007/08, 2008/09, and 2009/10) and demographic characteristics (2007/08 and 2008/09), Kentucky Department of Education Commonwealth Accountability Testing System results for KCCT (2006/07 and 2007/08) and EXPLORE (2006/07 and 2007/08), and data provided directly by schools (2007/08 and 2008/09).

Table H4. Sensitivity analysis of exploratory impact findings for student achievement results in grade 10 pre-algebra/algebra on the PLAN assessment, by cohort

Type of sensitivity analysis	Student subgroup	Treatment group mean (standard deviation)	Control group mean (standard deviation)	Difference (standard error)	Estimated impact in effect-size units[a]	*p*
1a) PLAN outcome adjusted for pretest variables only	Cohort I	6.75 (2.87)	6.94 (3.06)	–0.19	–0.07	
	Cohort II	6.56 (2.47)	6.95 (2.76)	–0.39	–0.13	
	Difference	–2.68	–2.91	0.20 (0.36)	0.07	0.58
1b) PLAN outcome adjusted for KCCT pretest variables only	Cohort I	6.65 (2.87)	6.98 (3.06)	–0.33	–0.12	
	Cohort II	6.70 (2.47)	6.95 (2.76)	–0.25	–0.09	
	Difference	–0.05 (2.68)	0.03 (2.91)	–0.08 (0.39)	–0.03	0.84
1c) PLAN outcome adjusted for EXPLORE pretest variables only	Cohort I	6.79 (2.87)	6.94 (3.06)	–0.15	–0.06	
	Cohort II	6.54 (2.47)	7.03 (2.76)	–0.49	–0.17	
	Difference	0.25 (2.68)	–0.09 (2.91)	0.34 (0.35)	0.12	0.34
2a) PLAN outcome adjusted for student and school characteristics *except* EXPLORE pretest	Cohort I	6.67 (2.87)	6.87 (3.06)	–0.20	–0.07	
	Cohort II	6.66 (2.47)	7.02 (2.76)	–0.36	–0.12	
	Difference	0.01 (2.68)	–0.15 (2.91)	0.16 (0.37)	0.06	0.66

Type of sensitivity analysis	Student subgroup	Treatment group mean (standard deviation)	Control group mean (standard deviation)	Difference (standard error)	Estimated impact in effect-size units[a]	*p*
2b) PLAN outcome adjusted for student and school characteristics *except* KCCT pretest	Cohort I	6.71 (2.87)	6.82 (3.06)	–0.11	–0.04	
	Cohort II	6.63 (2.47)	7.06 (2.76)	–0.43	–0.15	
	Difference	0.08 (2.68)	–0.24 (2.91)	0.32 (0.40)	0.11	0.45
3) PLAN outcome adjusted for student and school characteristics, excluding students enrolled in a part-year or part-credit course	Cohort I	6.72 (2.87)	6.83 (3.06)	–0.11	–0.04	
	Cohort II	6.58 (2.47)	6.95 (2.76)	–0.37	–0.13	
	Difference	0.14 (2.68)	–0.12 (2.91)	0.26 (0.39)	0.09	0.50
4) PLAN outcome adjusted for student and school characteristics, excluding the six duplicate schools in cohort I that were rerandomized in cohort II	Cohort I	6.71 (2.87)	6.87 (3.06)	–0.16	–0.06	
	Cohort II	6.64 (2.47)	6.94 (2.76)	0.30	0.10	
	Difference	0.07 (2.68)	–0.07 (2.91)	0.14 (0.43)	0.05	0.74

a. Effect sizes are calculated using Cohen's d with the pooled standard deviation.
* Difference (of estimated intent-to-treat impact) is statistically significant at the 95 percent confidence level using a two-tailed test.
Note: Means are regression adjusted using the same set of baseline covariates as the confirmatory analysis. Valid *n* for PLAN = 5,863 (cohort I = 2,679 and cohort II = 3,184). The *p*-value is for the coefficient on the interaction term between treatment and the cohort I dummy variable.
Source: Authors' analysis based on Kentucky Department of Education data for enrollment (2007/08, 2008/09, and 2009/10) and demographic characteristics (2007/08 and 2008/09), Kentucky Department of Education Commonwealth Accountability Testing System results for KCCT (2006/07 and 2007/08) and EXPLORE (2006/07 and 2007/08), and data provided directly by schools (2007/08 and 2008/09).

Table H5. Sensitivity analysis of exploratory impact findings for student achievement results in grade 10 pre-algebra/algebra on the PLAN assessment, by rural status

Type of sensitivity analysis	School subgroup	Treatment group mean (standard deviation)	Control group mean (standard deviation)	Difference (standard error)	Estimated impact in effect-size units[a]	*p*
1a) PLAN outcome adjusted for pretest variables only	Rural	6.78 (2.56)	6.76 (2.80)	0.02	0.01	
	Nonrural	6.52 (2.82)	7.50 (3.01)	–0.98	–0.37	
	Difference	0.26 (2.68)	–0.74 (2.91)	1.00* (0.35)	0.36	<0 .01
1b) PLAN outcome adjusted for KCCT pretest variables only	Rural	6.61 (2.56)	6.74 (2.80)	–0.13	–0.04	
	Nonrural	6.79 (2.82)	7.61 (3.01)	–0.82	–0.31	
	Difference	–0.18	–0.87	0.72	0.26	0.06

Type of sensitivity analysis	School subgroup	Treatment group mean (standard deviation)	Control group mean (standard deviation)	Difference (standard error)	Estimated impact in effect-size units[a]	*p*
		(2.68)	(2.91)	(0.37)		
1c) PLAN outcome adjusted for EXPLORE pretest variables only	Rural	6.86 (2.56)	6.81 (2.80)	0.05	0.02	
	Nonrural	6.45 (2.82)	7.52 (3.01)	–1.06	–0.41	
	Difference	0.41 (2.68)	–0.70 (2.91)	1.12* (0.34)	0.40	< .001
2a) PLAN outcome adjusted for student and school characteristics *except* EXPLORE pretest	Rural	6.66 (2.56)	6.74 (2.80)	–0.07	–0.02	
	Nonrural	6.75 (2.82)	7.35 (3.01)	–0.60	–0.23	
	Difference	–0.09 (2.68)	–0.61 (2.91)	0.50 (0.41)	0.18	0.22
2b) PLAN outcome adjusted for student and school characteristics *except* KCCT pretest	Rural	6.82 (2.56)	6.79 (2.80)	0.02	0.01	
	Nonrural	6.52 (2.82)	7.30 (3.01)	–0.78	–0.30	
	Difference	0.30 (2.68)	–0.50 (2.91)	0.80 (0.42)	0.29	0.06
3) PLAN outcome adjusted for student and school characteristics, excluding students enrolled in a part-year or part-credit course	Rural	6.77 (2.56)	6.68 (2.80)	0.09	0.03	
	Nonrural	6.51 (2.82)	7.35 (3.01)	–0.84	–0.32	
	Difference	0.26 (2.68)	–0.67 (2.91)	0.93* (0.41)	0.33	0.02
4) PLAN outcome adjusted for student and school characteristics, excluding the six duplicate schools in cohort I that were rerandomized in cohort II	Rural	6.78 (2.56)	6.69 (2.80)	0.09	0.03	
	Nonrural	6.60 (2.82)	7.32 (3.01)	–0.72	–0.28	
	Difference	0.17 (2.68)	–0.63 (2.91)	0.81 (0.43)	0.29	0.06

a. Standardized difference in terms of standard deviations of the study control group distribution.

* Difference (of estimated intent-to-treat impact) is statistically significant at the 95 percent confidence level using a two-tailed test.

Note: Means are regression adjusted using the same set of baseline covariates as the confirmatory analysis. Valid *n* for PLAN = 5,863 (rural = 3,631 and nonrural = 2,232). The *p*-value is for the coefficient on the interaction term between treatment and rural.

Source: Authors' analysis based on Kentucky Department of Education data for enrollment (2007/08 and 2008/09) and demographic characteristics (2007/08 and 2008/09), Kentucky Department of Education Commonwealth Accountability Testing System results for KCCT (2006/07 and 2007/08), EXPLORE (2006/07 and 2007/08), and PLAN (2007/08 and 2008/09); and data provided directly by schools (2007/08 and 2008/09).

Tables H6, H7, and H8 present the results from the sensitivity analysis for the exploratory impact analysis on grade 10 math course enrollment in the postintervention year.

Table H6. Sensitivity analysis of confirmatory impact findings for grade 10 math course enrollment, by gender

Type of sensitivity analysis	Student subgroup	Treatment group probability	Control group probability	Difference[a]	Odds ratio (standard error)	Log odds ratio	*p*
1a) Grade 10 math enrollment outcome adjusted for pretest variables only	Male	0.84	0.83	0.01	1.07	0.07	
	Female	0.89	0.90	–0.01	0.86	–0.15	
	Difference	–0.05	–0.07	0.02	1.19 (0.20)	0.17	0.29
1b) Grade 10 math enrollment outcome adjusted for KCCT pretest variables only	Male	0.84	0.84	0.00	1.04	0.04	
	Female	0.89	0.90	–0.01	0.87	–0.14	
	Difference	–0.05	–0.06	0.01	1.20 (0.20)	0.18	0.27
1c) Grade 10 math enrollment outcome adjusted for EXPLORE pretest variables only	Male	0.84	0.84	0.00	1.04	0.04	
	Female	0.89	0.90	–0.01	0.86	–0.15	
	Difference	–0.05	–0.06	0.01	1.15 (0.19)	0.14	0.38
2a) Grade 10 math enrollment outcome adjusted for student and school characteristics *except* EXPLORE pretest	Male	0.84	0.83	0.01	1.03	0.03	
	Female	0.88	0.90	–0.02	0.86	–0.16	
	Difference	–0.05	–0.07	0.02	1.21 (0.20)	0.19	0.25
2b) Grade 10 math enrollment outcome adjusted for student and school characteristics *except* KCCT pretest	Male	0.84	0.83	0.01	1.07	0.07	
	Female	0.88	0.89	–0.01	0.90	–0.11	
	Difference	–0.05	–0.07	0.02	1.18 (0.19)	0.17	0.33
3) Grade 10 math enrollment outcome adjusted for student and school characteristics, excluding students enrolled in a part-year or part-credit course	Male	0.84	0.83	0.01	1.09	0.09	
	Female	0.88	0.90	–0.02	0.85	–0.17	
	Difference	–0.04	–0.07	0.03	1.41 (0.25)	0.34	0.06
4) Grade 10 math enrollment outcome adjusted for student and school characteristics, excluding the six duplicate schools in cohort I that were rerandomized in cohort II	Male	0.83	0.83	0.00	1.00	0.00	
	Female	0.88	0.89	–0.01	0.87	–0.13	
	Difference	–0.05	–0.06	0.01	1.12 (0.19)	0.11	0.51

a. Impact estimates are estimated using the regression equation to calculate each student's probability of enrolling in a higher-level math course using individual student characteristics, and then generating group-level means for the treatment and control groups using the treatment variable. The impact estimate is the difference-in-difference, which is the difference between males in the treatment and control groups minus the difference between females in the treatment and control groups.

Note: Predicted probabilities are regression adjusted using the same set of baseline covariates as the confirmatory analysis. Valid *n* for grade 10 math course enrollment = 6,409 (male = 3,311, female = 3,084, and missing gender = 14). The *p*-value is for the coefficient on the interaction term between treatment and male.

Source: Authors' analysis based on Kentucky Department of Education data for enrollment (2007/08, 2008/09, and 2009/10) and demographic characteristics (2007/08 and 2008/09), Kentucky Department of Education Commonwealth Accountability Testing System results for KCCT (2006/07 and 2007/08) and EXPLORE (2006/07 and 2007/08), and data provided directly by schools (2007/08 and 2008/09).

Table H7. Sensitivity analysis of confirmatory impact findings for grade 10 math course enrollment, by cohort

Type of sensitivity analysis	Student subgroup	Treatment group probability	Control group probability	Difference[a]	Odds ratio (standard error)	Log odds ratio	*p*
1a) Grade 10 math enrollment outcome adjusted for pretest variables only	Cohort I	0.87	0.87	0.00	1.01	0.01	
	Cohort II	0.86	0.87	–0.01	0.93	–0.07	
	Difference	0.01	0.01	0.01	1.20 (0.82)	0.19	0.79
1b) Grade 10 math enrollment outcome adjusted for KCCT pretest variables only	Cohort I	0.87	0.87	0.00	1.01	0.01	
	Cohort II	0.86	0.87	–0.01	0.93	–0.07	
	Difference	0.01	0.01	0.01	1.16 (0.76)	0.15	0.82
1c) Grade 10 math enrollment outcome adjusted for EXPLORE pretest variables only	Cohort I	0.87	0.87	0.00	1.01	0.01	
	Cohort II	0.86	0.87	–0.01	0.93	–0.07	
	Difference	0.01	0.00	0.01	1.34 (0.87)	0.30	0.65
2a) Grade 10 math enrollment outcome adjusted for student and school characteristics *except* EXPLORE pretest	Cohort I	0.88	0.86	0.01	1.12	0.11	
	Cohort II	0.85	0.87	–0.02	0.84	–0.18	
	Difference	0.03	0.00	0.03	0.89 (0.59)	–0.11	0.86
2b) Grade 10 math enrollment outcome adjusted for student and school characteristics *except* KCCT pretest	Cohort I	0.87	0.86	0.02	1.14	0.13	
	Cohort II	0.84	0.86	–0.02	0.89	–0.12	
	Difference	0.03	0.00		0.94 (0.61)	–0.06	0.93
3) Grade 10 math enrollment outcome adjusted for student and school characteristics, excluding students enrolled in a part-year or part-credit course	Cohort I	0.87	0.86	0.01	1.10	0.10	
	Cohort II	0.85	0.86	–0.01	0.89	–0.12	
	Difference	0.03	0.00	0.03	0.81 (0.59)	–0.21	0.77
4) Grade 10 math enrollment outcome adjusted for student and school characteristics,	Cohort I	0.87	0.86	0.01	1.12	0.11	
	Cohort II	0.83	0.86	–0.03	0.78	–0.25	
	Difference	0.05	0.00	0.05	1.35 (0.99)	0.30	0.68

Type of sensitivity analysis	Student subgroup	Treatment group probability	Control group probability	Difference[a]	Odds ratio (standard error)	Log odds ratio	*p*
excluding the six duplicate schools in cohort I that were rerandomized in cohort II							

a. Impact estimates are estimated using the regression equation to calculate each student's probability of enrolling in a higher-level math course using individual student characteristics, and then generating group-level means for the treatment and control groups using the treatment variable. The impact estimate is the difference-in-difference, which is the difference between cohort I students in the treatment and control groups minus the difference between cohort II students in the treatment and control groups.

Note: Predicted probabilities are regression adjusted using the same set of baseline covariates as the confirmatory analysis. Valid *n* for grade 10 math course enrollment = 6,409 (cohort I = 3,437 and cohort II = 2,972). The *p*-value is for the coefficient on the interaction term between treatment and the cohort I dummy variable.

Source: Authors' analysis based on Kentucky Department of Education data for enrollment (2007/08, 2008/09, and 2009/10) and demographic characteristics (2007/08 and 2008/09), Kentucky Department of Education Commonwealth Accountability Testing System results for KCCT (2006/07 and 2007/08) and EXPLORE (2006/07 and 2007/08), and data provided directly by schools (2007/08 and 2008/09).

Table H8. Sensitivity analysis of confirmatory impact findings for grade 10 math course enrollment for school subgroups, by rural status

Type of sensitivity analysis	School subgroup	Treatment group probability	Control group probability	Difference[a]	Odds ratio (standard error)	Log odds ratio	*p*
1a) Grade 10 math enrollment outcome adjusted for pretest variables only	Rural	0.91	0.88	0.03	1.27	0.24	
	Nonrural	0.82	0.82	0.00	0.98	–0.02	
	Difference	0.09	0.06	0.03	1.71 (1.16)	0.54	0.43
1b) Grade 10 math enrollment outcome adjusted for KCCT pretest variables only	Rural	0.91	0.89	0.02	1.29	0.25	
	Nonrural	0.81	0.83	–0.02	0.87	–0.13	
	Difference	0.10	0.05	0.05	1.40 (0.94)	0.34	0.62
1c) Grade 10 math enrollment outcome adjusted for EXPLORE pretest variables only	Rural	0.91	0.88	0.03	1.30	0.26	
	Nonrural	0.82	0.83	–0.01	0.95	–0.06	
	Difference	0.09	0.06	0.03	1.63 (1.03)	0.49	0.44
2a) Grade 10 math enrollment outcome adjusted for student and school characteristics *except* EXPLORE pretest	Rural	0.90	0.88	0.02	1.23	0.21	
	Nonrural	0.80	0.83	–0.03	0.85	–0.16	
	Difference	0.10	0.06	0.04	1.32 (0.95)	0.28	0.71
2b) Grade 10 math enrollment outcome adjusted for student and school characteristics *except* KCCT pretest	Rural	0.90	0.88	0.02	1.24	0.22	
	Nonrural	0.81	0.82	–0.01	0.89	–0.11	
	Difference	0.10	0.06	0.04	1.45 (0.99)	0.37	0.59
3) Grade 10 math enrollment outcome adjusted for student and school characteristics, excluding students	Rural	0.90	0.89	0.01	1.15	0.14	
	Nonrural	0.79	0.82	–0.03	0.84	–0.18	
	Difference	0.11	0.07	0.04	1.23 (0.96)	0.21	0.80

Type of sensitivity analysis	School subgroup	Treatment group probability	Control group probability	Difference[a]	Odds ratio (standard error)	Log odds ratio	*p*
enrolled in a part-year or part-credit course							
4) Grade 10 math enrollment outcome adjusted for student and school characteristics, excluding the six duplicate schools in cohort I that were rerandomized in cohort II	Rural	0.89	0.88	0.01	1.01	0.01	
	Nonrural	0.82	0.82	0.00	0.99	–0.01	
	Difference	0.07	0.06	0.01	1.05 (0.81)	0.05	0.95

a. Impact estimates are estimated using the regression equation to calculate each student's probability of enrolling in a higher-level math course using individual student characteristics, and then generating group-level means for the treatment and control groups using the treatment variable. The impact estimate is the difference-in-difference, which is the difference between rural students in the treatment and control groups minus the difference between nonrural students in the treatment and control groups.

Note: For each subgroup, the predicted probabilities are regression adjusted using the same set of baseline covariates as the confirmatory analysis. Valid n for grade 10 math course enrollment = 6,409 (rural = 3,934 and nonrural = 2,475). The p-value is for the coefficient on the interaction term between treatment and rural.

Source: Authors' analysis based on Kentucky Department of Education data for enrollment (2007/08, 2008/09, and 2009/10) and demographic characteristics (2007/08 and 2008/09), Kentucky Department of Education Commonwealth Accountability Testing System results for KCCT (2006/07 and 2007/08) and EXPLORE (2006/07 and 2007/08), and data provided directly by schools (2007/08 and 2008/09).

References

ACT (American College Testing). (2007a). *PLAN technical manual*. Iowa City, IA: Author. Retrieved June 9, 2008, from http://www.act.org/plan/pdf/PlanTechnicalManual.pdf

ACT. (2007b). *EXPLORE technical manual*. Iowa City, IA: Author. Retrieved June 9, 2008, from http://www.act.org/explore/pdf/TechManual.pdf

ACT. (2009). *ACT's PLAN assessment: Components*. Iowa City, IA: Author. Retrieved November 21, 2009, from http://www.act.org/plan/components.html

Adelman, C. (1999). *Answers in the tool box: Academic intensity, attendance patterns, and bachelor's degree attainment.* Washington, DC: U.S. Department of Education.

Adelman, C. (2006). *The toolbox revisited: Paths to degree completion from high school to college.* Washington, DC: United States Department of Education, Office for Educational Research and Improvement.

Altman, D. G., Schulz, K. F., Moher, D., Egger, M., Davidoff, F., Elbourne, D., et al. (2001). The revised CONSORT statement for reporting randomized trails: Explanation and elaboration. *Annals of Internal Medicine, 134*(8), 663–694.

Banilower, E., Heck, D., & Weiss, I. (2005). Can professional development make the vision of the standards a reality? The impact of the National Science Foundation's local systemic change through teacher enhancement initiative. *Journal of Research in Science Teaching*, *44*(3), 375–395.

Banilower, E., & Shimkus, E. (2004). *Professional development observation study.* Chapel Hill, NC: Horizon Research.

Barrow, L., Markman, L., & Rouse, C. E. (2009). Technology's edge: The educational benefits of computer-aided instruction. *American Economic Journal: Economic Policy*, *1*(1), 52–74.

Billmeyer, R., & Barton, M. L. (1998). *Teaching reading in the content areas, if not me then who?* Aurora, CO: Mid-continent Regional Educational Laboratory.

Bloom, H. S., Hill, C. J., Black, A. R, & Lipsey, M. W. (2008). *Performance trajectories and performance gaps as achievement effect-size benchmarks for educational interventions.* New York: MDRC. Retrieved March 9, 2009, from http://www.mdrc.org/publications/500/full.pdf

Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher, 33*(8), 3–15.

Campuzano, L., Dynarski, M., Agodini, R., & Rall, K. (2009). *Effectiveness of reading and mathematics software products: Findings from two student cohorts.* (NCEE 2009-4041). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. Retrieved March 29, 2010 from http://ies.ed.gov/ncee/pubs/20094041/pdf/20094041.pdf

Carey, N., & Frechtling, J. (1997). *Best practice in action: Follow-up survey on teacher enhancement programs.* Arlington, VA: National Science Foundation.

Carr, M., & Davis, H. (2001). Gender differences in arithmetic strategy use: A function of skill and preference. *Contemporary Educational Psychology*, *26*(3), 330–347.

CNA. (2005a). *RAC report to the U.S. Dept. of Education*. Alexandria, VA: The CNA Corporation, Appalachia Regional Advisory Committee.

CNA. (2005b). *RAC report to the U.S. Dept. of Education, Synthesis*. Alexandria, VA: The CNA Corporation, Appalachia Regional Advisory Committee.

Cohen, D. K. (1990). A revolution in one classroom: The case of Mrs. Oublier. *Educational Evaluation and Policy Analysis, 12*(3), 311–329.

Cohen, D. K., & Hill, H. C. (2001). *Learning policy: When state education reform works.* New Haven, CT: Yale University Press.

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlational analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cooper, H., Nye, B., Charlton, K., Lindsay, J., & Greathouse, S. (1996). The effects of summer vacation on achievement test scores: A narrative and meta-analytic review. *Review of Educational Research, 66*(3), 227–268.

Cox, D. R. (1970). *Analysis of binary data*. New York: Chapman & Hall/CRC.

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. New York: Harcourt Brace Jovanovich College Publishers.

Darling-Hammond, L. (1997). *The right to learn: A blueprint for creating schools that work*. San Francisco: Jossey-Bass.

Desimone, L. M. (2003). Toward a more refined theory of school effects: A study of the relationship between professional community and mathematic teaching in early elementary school. In C. Miskel and W. Hoy (Eds.), *Research and theory in educational administration* 87-134. Greenwich, CT: Information Age.

Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualization and measures. *Educational Researcher*, *38*(3), 181–199.

Dynarski, M., Agodini, R., Heaviside, S., Novak, T., Carey, T., and Campuzano, L., et al. (2007). *Effectiveness of reading and mathematics software products: Findings from the first student cohort.* Report to Congress (NCEE 2007-4005). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. Retrieved April 19, 2008 from http://ies.ed.gov/ncee/pdf/20074005.pdf

Educational Testing Service. (2004). *Administrator's manual for the Assessment of Algebraic Understanding*. Princeton, NJ: Educational Testing Service.

Friedman, L. (1995). The space factor in mathematics: Gender differences. *Review of Educational Research, 65*(1), 22–50.

Fullan, M. (1991). *The new meaning of educational change.* New York: Teachers College Press.

Garet, M., Porter, A., Desimone, L., Birman, B., & Yoon, K. (2001). What makes professional development effective? Results from a national sample of teachers. *American Education Research Journal, 38*(4), 915–945.

Garnham, C., & Kaleta, R. (2002). Introduction to hybrid courses. *Teaching with Technology Today, 8*(6). Retrieved April 1, 2008 from http://www.uwsa.edu/ttt/articles/garnham.htm

Geary, D. C., Saults, S. J., Liu, F., & Hoard, M. K. (1999). Sex differences in spatial cognition, computational fluency, and arithmetical reasoning. *Journal of Experimental Child Psychology, 77*(4), 337–353.

Ginsburg-Block, M. D., & Fantuzzo, J. W. (1998). An evaluation of the relative effectiveness of NCTM standards-based interventions for low-achieving urban elementary students. *Journal of Educational Psychology, 90*(3), 560–569.

Guskey, T. R. (1994). Results-oriented professional development: In search of an optimal mix of effective practices. *Journal of Staff Development, 15*(4), 42–50.

Inan, F.A., Lowther, D.L., Ross, S.M., & Strahl, J.D. (2010). Pattern of Classroom Activities during Students Use of Computers: Relations between Instructional Strategies and Computer Applications. *Teaching and Teacher Education, 26*(3), 540-546.

ISTE (International Society for Technology in Education). (1998). *National educational technology standards for students, (NETS-S).* Eugene, OR: Author.

ISTE. (2000). National educational technology standards for teachers, (NETS-T). Eugene, OR: Author.

Kennedy, M. M. (1998). *Form and substance in in-service teacher education* (Research Monograph No. 13). Arlington, VA: National Science Foundation.

Kentucky Department of Education. (2006a). *Minimum high school graduation requirements*. Frankfort, KY. Retrieved February 11, 2010, from http://www.education.ky.gov/kde/instructional+resources/secondary+and+virtual+learning/high+school/minimum+high+school+graduation+requirements.htm

Kentucky Department of Education. (2006b). *The Program of Studies – Revised 2006: High School*. Retrieved February 11, 2010, from http://www.education.ky.gov/users/jwyatt/POS/High.doc

Kentucky Department of Education. (2009). *Commonwealth accountability testing system technical report, 2007-08*. Frankfort, KY. Retrieved April 19, 2009, from http://www.education.ky.gov/KDE/Administrative+Resources/Testing+and+Reporting+/Kentucky+School+Testing+System/Accountability+System/Technical+Manual+2008.htm

Kentucky Department of Education. (2010). *Highly skilled educators*. Frankfort, KY. Retrieved November 5, 2010 from

http://www.education.ky.gov/KDE/Administrative%2BResources/School%2BImprovement/Assistance%2Bto%2BSchools/Highly%2BSkilled%2BEducators/

Kirkpatrick, H., & Cuban, L. (1998). Computers make kids smarter — right? *Technos, 7*(2), 26–31.

Kulik, J. A. (2003). *Effects of using instructional technology in elementary and secondary schools: What controlled evaluation studies say*. Arlington, VA: SRI International.

Lewis, E. M., Ross, S. M., & Alberg, M. (1999). *School observation measure: Reliability analysis*. Memphis, TN: The University of Memphis, Center for Research in Educational Policy.

Lieberman, A. (1996). Creating intentional learning communities. *Educational Leadership*, *54*(3), 51–55.

Little, J. W. (1993). Teachers' professional development in a climate of educational reform. *Educational Evaluation and Policy Analysis, 15*(2), 129–151.

Loeb, S., & Reininger, M. (2004). *Public policy and teacher labor market: What we know and why it matters.* East Lansing, MI: Michigan State University, Education Policy Center.

Long, M. C., Iatarola, P., & Conger, D. (2009). Explaining gaps in readiness for college-level math: The role of high school courses. *Education Finance and Policy, 4*(1), 1–33.

Loucks-Horsley, S., Hewson, P., Love, N., & Stiles, K. (1998). *Designing professional development for teachers of science and mathematics*. Thousand Oaks, CA: Corwin Press.

Lowther, D. L. (2006). *Algebra I Quality Assessment (AQA).* Memphis, TN: The University of Memphis, Center for Research in Educational Policy.

Lowther, D. L.,& Ross, S. M. (2000). *Teacher Technology Questionnaire (TTQ).* Memphis, TN: The University of Memphis, Center for Research in Educational Policy.

Lowther, D. L., & Ross, S. M. (2002). *Rubric for Student-Centered activities (RSCA).* Memphis, TN: The University of Memphis, Center for Research in Educational Policy.

Mayer, D. P. (1998). Do new teaching standards undermine performance on old tests? *Educational Evaluation and Policy Analysis, 20*(2), 53–73.

McCaffrey, D., Hamilton, L., Stecher, B., Klein, S., Bugliari, D., & Robyn, A. (2001). Interactions among instructional practices, curriculum, and student achievement: The case of standards-based high school mathematics. *Journal for Research in Mathematics Education, 32*(5), 493–517.

McClure, C., Redfield, D., & Hammer, P. (2003). *Recruiting and retaining high-quality teachers in rural areas*. Charleston, WV: Appalachia Educational Laboratory.

Monk, D.H. (2007). Recruiting and retaining high-quality teachers in rural areas. *The Future of Children, Excellence in the Classroom*, *17*(1), 155-174.

Moyer, P. S. (2001). Are we having fun yet? How teachers use manipulatives to teach mathematics. *Educational Studies in Mathematics, 47*(2), 175–197.

Muller, C., and Schiller, K. S. (2000). Leveling the playing field? Students' educational attainment and states' performance testing. *Sociology of Education, 73*(3), 196–218.

Murphy, R., Penuel, W., Means, B., Korbak, C., & Whaley, A. (2001). *E-DESK: A review of recent evidence on the effectiveness of discrete educational software*. Menlo Park, CA: SRI International.

NACOL (North American Council for Online Learning). (2007). *National standards of quality for online courses.* Retrieved February 11, 2010, from http://www.inacol.org/research/nationalstandards/NACOL%20Standards%20Quality%20Online%20Courses%202007.pdf

National Research Council. (1989). *Everybody counts: A report to the nation on the future of mathematics education*. Washington, DC: Author.

National Standards Development Council. (2001). *National standards for staff development,* rev. 2001. Retrieved October 5, 2008 from http://www.nsdc.org/standards/index.cfmNCTM (National Council of Teachers of Mathematics). (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: Author.

NCTM. (1991). *Professional standards for school mathematics*. Reston, VA: Author.

NCTM. (2000). *Principles and standards for school mathematics*. Reston, VA: Author.

Puma, M., Olsen, R., Bell, S., & Price, C. (2009). *What to do when data are missing in group randomized controlled trials* (NCEE 2009-0049). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Raudenbush, S. W., Spybrook, J., Liu, X. F., & Congdon, R. (2005). Optimal design for longitudinal and multilevel research (V1.55). [Computer software]

Rosenholtz, S. J. (1989). Workplace conditions that affect teacher quality and commitment: Implications for teacher induction programs. *Elementary School Journal, 89*(4), 421–439.

Ross, S. M., Smith, L. J., & Alberg, M. (1999). *The school observation measure (SOM ©)*. Memphis, TN: The University of Memphis, Center for Research in Educational Policy.

Ross, S.M., Smith, L.J., Alberg, M., & Lowther, D.L. (2004). *Using Classroom Observation as a Research and Formative Evaluation Tool in Educational Reform: The School Observation Measure* in H. Waxman (Ed.). Observation Research in U.S. Classrooms: New Approaches for Understanding Cultural and Linguistic Diversity (pp. 144-173). Cambridge, MA: Cambridge University Press.

Schochet, P. (2008). *Technical methods report: Guidelines for multiple testing in impact evaluations* (NCEE 2008-4018.). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Smith, T. M., Desimone, L. M., Zeidner, T., Dunn, A. C., Bhatt, M., & Rumyantseva, N. (2007). Inquiry-oriented instruction in science: Who teaches that way? *Educational Evaluation and Policy Analysis, 9*(29), 169–199.

Southern Region Education Board. (2001). *Spotlight on Algebra I, user's guide, version 5.0.* [Computer software]

Sterbinsky, A., & Burke, D. (2004). *Summary of CSRTQ reliability studies. Technical report.* Memphis, TN: The University of Memphis, Center for Research in Educational Policy.

Strauss, R. (1999). Who gets hired to teach? The case of Pennsylvania. In M. Kanstroom (Ed.), *Better Teachers, Better Schools,* 117-119. New York, NY: Fordham Foundation.

U.S. Department of Education, Institute of Education Science. (n.d.). What Works Clearinghouse. Retrieved May 15, 2010, from http://ies.ed.gov/ncee/wwc/references/iDocViewer/Doc.aspx?docId=9#dichotomous

Waxman, H. C., Lin, M.-F., & Michko, G. M. (2003). *A meta-analysis of the effectiveness of teaching and learning with technology on student outcomes.* Naperville, IL: Learning Point Associates, for the U.S. Department of Education, Institute of Education Sciences, North Central Regional Educational Laboratory. Retrieved April 19, 2008, from http://www.ncrel.org/tech/effects2/waxman.pdf

www.ed.gov ies.ed.gov